

# LARGE DEVIATIONS FOR SPATIALLY EXTENDED RANDOM NEURAL NETWORKS

TANGUY CABANA AND JONATHAN TOUBOUL

**ABSTRACT.** We investigate the asymptotic mesoscopic behavior of a spatially extended stochastic neural networks dynamics in random environment with highly random connectivity weights. These systems model the spatiotemporal activity of the brain, thus feature (i) communication delays depending on the distance between cells and (ii) heterogeneous synapses: connectivity coefficients are random variables whose law depends on the neurons positions and whose variance scales as the inverse of the network size. When the weights are independent Gaussian random variables, we show that the empirical measure satisfies a large-deviation principle. This holds under a technical condition on time horizon, noise and heterogeneity in the general case, and with no restriction in the case where delays do not depend on space. The associated good rate function achieves its minimum at a unique spatially extended probability measure, implying convergence of the empirical measure and propagation of chaos. The limit is characterized through complex non Markovian implicit equation in which the network interaction term is replaced by a non-local Gaussian process whose statistics depend on the solution over the whole neural field. We further demonstrate the universality of this limit, in the sense that neuronal networks with non-Gaussian interconnection weights converge towards it provided that synaptic weights have a sufficiently fast decay.

## CONTENTS

1. Introduction	2
1.1. Biological background	3
1.2. Microscopic Neuronal Network Model	4
2. Statement of the results	5
3. Large Deviation Principle	10
3.1. Construction of the good rate function	10
3.2. Upper-bound and Tightness	24
4. Identification of the mean-field equations	27
4.1. Variational characterization of the minima of $H$	27
4.2. Uniqueness of the minimum	29
4.3. Convergence of the process	35
5. Non Gaussian connectivity weights	36
Appendix A. A priori estimates for single neurons	46
Appendix B. Proof of lemma 2: regularity of the solutions for the limit equation	46
Appendix C. Proof of lemma 14: approximation of $\Gamma$ by $\Gamma_\nu$	50
Appendix D. Non-Gaussian estimates	53

---

*Date:* October 27, 2015.

## 1. INTRODUCTION

We study the asymptotic behavior of spatially-extended neuronal networks with heterogeneous interconnections at a mesoscopic scale in which averaging effects occur but where one can still resolve fine spatial structures. In detail, we will consider a stochastic network equation of size  $N$  in random environment, in which

- neurons have random locations on a compact set  $D \subset \mathbb{R}^d$ ,
- the amplitude of the interaction between two cells are heterogeneous. Their statistics depend on the cells positions and have a mean and a variance scaling as  $1/N$ ,
- neurons communicate after a delay, also depending on the cells locations, associated with transport and transmission of information.

Each network is characterized by a random *configuration* that does not evolve in time. Within this fixed network configuration, the state of each neuron is described by a stochastic nonlinear process. The motivation for developing this model lies in the understanding of spatio-temporal patterns of activity of the cortex, as we review in section 1.1. For spatially extended networks with “weak” interaction heterogeneities (variance of interconnection weights scaled as  $1/N^2$ ), coupling methods have been used to derive a non-local McKean-Vlasov thermodynamic limit [44, 43] where the effective interaction term involves a non-local integral in space. This limit depends explicitly on the averaged spatial structure of the brain, thus preserving important information on spatiotemporal patterns of activity [42]. In the context of networks on lattices with non-random synapses and no delay, compactness methods were used to show a convergence result towards a nonlinear Fokker-Planck equation [29]. For interacting heterogeneous diffusions with non-random interconnections, large-deviations techniques were developed [17] and convergence of double-layer empirical distributions including state variance and heterogeneity was proved. In all these cases, the heterogeneity of the interconnections was not sufficient to affect the asymptotic behavior.

Strongly stochastic synapses have been the object of intense studies in the domain of mathematical physics. Sophisticated techniques were developed in the context of spin glasses (see e.g. the reference books [40, 41]). Of particular relevance to our purposes, and in the same context, large-deviations techniques were devised for randomly connected networks with strongly heterogeneous interconnections [5, 25, 4, 3]. The methods were then adapted for biological neural networks in discrete time settings in a number of models [18, 15, 23], and were recently extended to continuous-time diffusions with multiple populations and delays [13]. In all these contributions, synaptic weights were considered Gaussian and the limit found involved an implicit effective interaction term that has a Gaussian law. Although methods of proof use Gaussian calculus, the Gaussian nature of the limit process does not require weights to be Gaussian, but similarly to the central limit theorem, is valid for a broad class of couplings. This was rigorously addressed in [32] in the case of discrete-time dynamics for weights with sub-Gaussian tails. It proves surprisingly complex to generalize their approach in a continuous-time setting.

We undertake in this manuscript the characterization of spatially extended networks with continuous-time dynamics and strongly heterogeneous synapses as motivated by the study of the spatio-temporal cortical patterns of activity. To this purpose, we combine large-deviations estimates and the methods developed for spatially extended particle systems to demonstrate the thermodynamic convergence of the network equation and identify their non-Markovian limit, for Gaussian and non-Gaussian synaptic weights. Before we proceed to the exposition of the setting and main results, we briefly review our motivation and model.

**1.1. Biological background.** It has been widely shown that mammalian brain displays precise spatiotemporal patterns of activity that correlate with brain states and cognitive processes. Classical examples include transient and local activation of specific regions in the cortex while recalling a memory (see e.g. [24]), visual illusions [27] or the propagation of a localized stimulus [33]. A popular and very efficient approach to describe these phenomena is the Wilson and Cowan neural field equation [45, 46], characterizing the spatiotemporal evolution of the activity  $u(r, t)$  of cells at location  $r$  on the neural field  $D$  through a simple integro-differential equation of type:

$$(1) \quad \frac{\partial u}{\partial t} = -u(r, t) + \int_D J(r, r') S(u(r', t)) dr' + I(r, t)$$

where  $I(r, t)$  represents the input to the population at location  $r$ ,  $J(r, r')$  is the averaged interconnection weight from neurons at location  $r'$  onto neurons at location  $r$  and the non-decreasing map  $S$  associates to a level of activity  $u$  the resulting spiking rate. This equation has been very successful in reproducing a number of biological phenomena, in particular working memory [28] and visual hallucination patterns [21, 8]. However, randomness is not explicitly present in it, and the relationship between the dynamics of individual cells and this macroscopic equation - a central problem in neuroscience [7] - is still elusive.

The present paper pursues the endeavor of addressing rigorously this relationship. Beyond its mathematical interest, this approach would provide a way to understand, from the biological viewpoint, the importance of individual cells or synaptic properties on brain's emerging behaviors. Specifically, this would provide a way to characterize the role of noise and heterogeneity, that were reported to be related to pathologies such as febrile seizures [2]. From the phenomenological viewpoint, it has also been shown that the variance of the weights can notably affect the network behavior and lead to phase transitions from trivial to chaotic solutions [37] or synchronization in two-populations networks [26].

The question of characterizing limits of large-scale dynamics of neuronal networks has a long history in neuroscience, and several mathematical and statistical physics methods were introduced. These range from PDE formalisms and kinetic equations [14, 34, 35] with deep applications to the visual system, moment reductions and master equations [30, 6], but also the development of specific Markov chain models reproducing in the thermodynamics limit the dynamics of Wilson-Cowan systems [9, 10, 6, 11, 12]. These techniques were generally developed in order to obtain limits of interconnected neurons through weakly stochastic synapses (typically constant or independent identically distributed synaptic weights with variance  $1/N^2$ , with  $N$  the typical number of incoming connections), and do not hold in the

case of *strongly stochastic synapses* whereby synaptic weights have a variance scaled by  $1/N$ .

**1.2. Microscopic Neuronal Network Model.** The macroscopic activity of cells relies on the collective activity emerging from a large number  $N$  of neurons that are distributed over the cortex, seen as a  $d$ -dimensional compact set  $D \subset \mathbb{R}^d$  ( $d$  is generally considered to be equal to 2, sometimes 3). The location of neuron  $i \in \{1, \dots, N\}$  is denoted  $r_i \in D$ , and we assume that locations are independently drawn according to a probability measure  $\pi \in \mathcal{M}_1^+(D)$  representing the density of neurons on the cortex, and assumed to be absolutely continuous with respect to Lebesgue's measure. The state of neuron  $i$  is described by a variable  $X^{i,N} \in \mathbb{R}^s$ , and we will assume here for simplicity that  $X^{i,N}$  is a scalar variable representing the voltage of each neuron and satisfying the equation:

$$(2) \quad dX_t^{i,N} = \left( f(r_i, t, X_t^{i,N}) + \sum_{j=1}^N J_{ij} S(X_{t-\tau_{ij}}^{j,N}) \right) dt + \lambda(r_i) dW_t^i,$$

where the map  $f(r, t, x)$  describes the intrinsic dynamics of a neuron at location  $r$ , time  $t$  and state  $x$ ,  $\lambda(r)$  the level of noise at location  $r$ , and where we assumed each neuron to be driven by an independent Brownian motions  $W_t^i$ . The interactions between cells are assumed, as in the classical firing-rate formalism [45, 46, 1], to be proportional to a sigmoidal transformation of their membrane potential  $S(X_t^{j,N})$ .  $S$  is a smooth (at least continuously differentiable) increasing map tending to 0 at  $-\infty$  and to 1 at  $\infty$ . The synaptic weight  $J_{ij}$  represents the amplitude and excitatory or inhibitory nature of the interaction depending on whether  $J_{ij} > 0$  or  $J_{ij} < 0$ . There is no connection between  $j$  and  $i$  when  $J_{ij} = 0$ . The parameters  $\tau_{ij}$  represent the delay of communication between the two neurons, and is assumed to be equal to a deterministic function of the location of neuron  $i$  and  $j$ :  $\tau_{ij} = \tau(r_i, r_j)$  (generally an affine function of the distance between cells  $\|r_i - r_j\|$  when spikes are assumed to be transmitted at constant speed).

The spatio-temporal activity of the cortex is obtained as a mesoscopic limit of cells activity that resolves distinct locations on the cortex, but where averaging effects related to the large dimension of the network are taken into account. In order to characterize these averaging effects, we will investigate the limit of the network as its size diverges. We thus need to describe how synaptic weights scale with the network size. Consistently with the underlying biological problem (see e.g. [43] and references therein), we assume that the connectivity weights  $J_{ij}$  are random variables whose law depends on the location of cells  $i$  and  $j$ , with mean  $J(r_i, r_j)/N$  and variance  $\sigma^2(r_i, r_j)/N$ . The scaling of the mean ensures that the interaction term does not diverge, while the scaling on the variance, slower than usual cases in  $1/N^2$  [39, 43] preserves a non-trivial contribution of the heterogeneous nature of the synaptic weights. Note that biologically, the synaptic weights cannot reach arbitrarily large or small values.

Before we proceed to rigorous developments, let us start by describing heuristically the large  $N$  behavior. One can generally get an intuition of the limit of such interacting systems by considering that the  $(X^{j,N}, r_j)$  are iid and independent of the connectivity matrix, and that the network equation converges towards a spatially extended process with law  $(\bar{X}(r), r)$ . This assumption is known as the Boltzman's "molecular chaos" (*Stoßzahlansatz*) hypothesis. Under these assumptions, one can

formally make the conjecture that the network interaction term  $\sum_{j=1}^N J_{ij} S(X_t^{j,N})$  converges, by virtue of a functional central limit theorem, towards a Gaussian process  $U_t^{\bar{X}}(r)$  with mean and covariance that are non-local (i.e. depending on the process  $X$  at all other locations), given by:

$$(3) \quad \begin{cases} \int_D J(r_i, r') \mathbb{E}[S(\bar{X}_t(r'))] d\pi(r') \\ \int_D \sigma(r, r')^2 \mathbb{E}[S(\bar{X}_{t-\tau(r, r')}(r')) S(\bar{X}_{s-\tau(r, r')}(r'))] d\pi(r') \end{cases}$$

and one thus obtains the implicit equation on  $\bar{X}$ :

$$(4) \quad d\bar{X}_t(r) = \left( f(r, t, \bar{X}_t(r)) + U_t^{\bar{X}}(r) \right) dt + \lambda(r) dW_t(r).$$

Interestingly, we recover an interaction term whose mean is exactly of the Wilson-Cowan type (1). Moreover, when  $f(r, t, x) = -x$ , solutions are Gaussian and their mean satisfies a Wilson-Cowan equation (1), in which the sigmoid function depends dynamically on noise and heterogeneity. In particular, we will see that Boltzmann's molecular chaos asymptotically occurs for any finite set of neurons. Note that Boltzmann's Stoßzahlansatz could indicate a certain degree of universality for the limit, as is displayed by the central limit theorem. In particular, it is possible that this limit remains valid for synaptic weights with bounded second moment. Universality will be partially addressed here, as we will prove the validity of the limit for sub-Gaussian-tailed synaptic weights, as well as for Gaussian-tailed synaptic weights under a short-time hypothesis.

The organization of the paper is as follows. We provide in section 2 the notations and main assumptions on the model, as well as a summary of the main results. Sections 3.1 and 4 deal with the case of Gaussian synaptic weights, and are respectively dedicated to the demonstration of a large deviations principle and to the identification of the limit. Section 5 is devoted to showing a general convergence result in the case where the synaptic weights are non-Gaussian, including in particular the biologically relevant case of bounded coefficients.

## 2. STATEMENT OF THE RESULTS

We investigate the thermodynamic limit of the neuronal network equations (2). These equations are diffusions in random environment, and thus exhibit two distinct sources of randomness:

- *Random environment:* the locations of neurons  $\mathbf{r} = (r_i)_{i \leq N}$  and synaptic weights  $J = (J_{ij})_{i, j \leq N}$  are random variables of a probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathcal{P})$ . They define the structure of the network, and are independent of the time-fluctuation of the states variable.
- *Stochastic dynamics:* states of neurons are stochastic variables, solutions of a SDE driven by a collection of independent  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ -Brownian motions  $(W_t^i)_{i \in \mathbb{N}}$ .

The dynamic of the  $X^i$  thus depends both on the random environment (i.e., the realization of locations  $\mathbf{r}$  and weights  $J$ ) and noise (the realization of the Brownian motions). We will denote by  $\mathcal{E}$  the expectation over the environment (i.e. with respect to the probability distribution  $\mathcal{P}$ ) and introduce the shorthand notation  $\mathcal{P}_J$  and  $\mathcal{E}_J$  the probability and expectation over the synaptic weights matrix  $J$  only (that is,  $\mathcal{P}$  and  $\mathcal{E}$  conditioned over the positions  $\mathbf{r}$ ). We recall that  $J$  depends on  $\mathbf{r}$ , but that the inverse is not true.

We work under a few regularity assumptions. In particular, we assume that the law of the synaptic weights is continuous in space. In details, although synaptic weights  $J_{ij}$  and  $J_{i'j'}$  are independent for  $i \neq i'$  or  $j \neq j'$ , we assume that their probability distribution continuously depends on the spatial location of the cells, in the sense that one can find a version  $\tilde{J}_{i'j'}$  of  $J_{i'j'}$  such that:

$$(5) \quad \mathcal{E}_J(|\tilde{J}_{i'j'} - J_{ij}|) \leq \frac{C}{N}(|r_i - r_{i'}| + |r_j - r_{j'}|).$$

for some  $C > 0$  independent of the neurons locations. Moreover, the dynamics of the neurons is assumed to satisfy the following assumptions:

- (1) The function  $f$  is  $K_f$ -Lipschitz continuous in its three variables.
- (2) The mean and variance of the weights  $J$  and  $\sigma$  are bounded and, respectively,  $K_J$  and  $K_\sigma$ -Lipschitz continuous in their second variable. We denote

$$\|J\|_\infty = \sup_{(r,r') \in D^2} |J(r,r')|, \quad \|\sigma\|_\infty = \sup_{(r,r') \in D^2} \sigma(r,r').$$

- (3)  $\tau : D^2 \rightarrow \mathbb{R}^+$  is Lipschitz continuous, with constant  $K_\tau$ . It is in particular bounded, by compactity of  $D$ . We denote by  $\tau$  its supremum.
- (4) The diffusion coefficient  $\lambda : D \rightarrow \mathbb{R}_+^*$  is a  $K_\lambda$  Lipschitz continuous and uniformly lowerbounded:  $\forall r \in D, \lambda(r) \geq \lambda_* > 0$ .

Let  $\mathcal{C}_\tau := \mathcal{C}([-\tau, 0], \mathbb{R})$ , and  $\mu_0 : D \rightarrow \mathcal{M}_1^+(\mathcal{C}_\tau)$  be an initial probability distribution mapping that depends continuously on  $r \in D$ . Throughout the paper, we consider that the network's initial conditions are independent realizations of  $\mu_0$ :

$$(6) \quad \text{Law of } (x_t)_{t \in [-\tau, 0]} = \bigotimes_{i=1}^N \mu_0(r_i).$$

It will often be useful to grant the existence of exponential quadratic moments to the solutions, and thus we will make the assumption that initial condition has the following moments condition:

$$(7) \quad \exists v > 0, \sup_{r \in D} \left\{ \int_{\mathcal{C}([-\tau, 0], \mathbb{R})} \exp \left\{ v \sup_{s \in [-\tau, 0]} |x_s|^2 \right\} d(\mu_0(r))(x) \right\} < \infty.$$

We further assume that their trajectories have the same regularity in time as the Brownian motion. The first question that may arise at this point is the well-posedness of the network system. Since the network equations constitute a standard delayed stochastic differential equation in dimension  $N$  with Lipschitz continuous drift and diffusion functions with linear growth property, standard theory on delayed stochastic differential equations [16, 31] ensures existence, uniqueness and square integrability of solutions:

**Proposition 1.** *For each  $\mathbf{r} \in D^N$ , and  $J \in \mathbb{R}^{N \times N}$  and  $T > 0$ , there exists a unique weak solution to the system (2) defined on  $[-\tau, T]$  with initial condition (6). Moreover, this solution is square integrable.*

*Remark 1.* Note that if the initial condition was given by  $(X_t^{i,N})_{t \in [-\tau, 0]} = \zeta^i$  with  $\zeta^i \stackrel{\mathcal{L}}{=} \mu_0(r_i)$ , we can of course prove strong existence and uniqueness of solutions.

We now work with an arbitrary fixed time  $T > 0$  and denote by  $Q_{\mathbf{r}}^N(J)$  the unique law solution of the network equations restricted to the  $\sigma$ -algebra  $\sigma(X_s^{i,N}, 1 \leq i \leq N, -\tau \leq s \leq T)$ .  $Q_{\mathbf{r}}^N(J)$  is a probability measure on  $\mathcal{C}^N$  where  $\mathcal{C}$  is the space of real

valued continuous functions of  $[-\tau, T]$ . This measure depends on the realizations of both the connectivity matrix  $J$ , and the locations of neurons  $\mathbf{r}$ . In order to characterize the behavior of the system as the network size diverges, we will show a Large Deviations Principle (LDP) for the empirical measure. This requires delicate estimates, combining different elements.

First, Sanov theorem states that, for independent copies of the same law  $\mu$  on a Polish space  $\Sigma$ , the empirical measure satisfies a full LDP with good rate function corresponding to the relative entropy  $I(\cdot|\mu)$  defined, for  $\nu \in \mathcal{M}_1^+(\Sigma)$ , by:

$$I(\nu|\mu) := \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu & \text{if } \nu \ll \mu, \\ \infty & \text{otherwise.} \end{cases}$$

Because of the connections, it is clear that Sanov's theorem does not apply as the states of neurons are not independent. Moreover, symmetry between cells is also broken by the choice of a realization of the interaction matrix. This motivates us to introduce the system without interaction. When neurons are not coupled (i.e.  $J_{ij} = 0$  for all  $(i, j)$ ), and locations are known, the law of neurons in position  $r \in D$  is given by the unique solution  $P_r$  of the one-dimensional standard SDE:

$$(8) \quad \begin{cases} dX_t = f(r, t, X_t)dt + \lambda(r)dW_t \\ (X_t)_{t \in [-\tau, 0]} \stackrel{\mathcal{L}}{=} \mu_0(r). \end{cases}$$

We denote by  $P_r$  the law of this process restricted to the  $\sigma$ -algebra  $\mathcal{G}_T = \sigma(X_s, s \leq T)$ ; it is a probability measure on the space  $\mathcal{C}$ . Remark that, by a direct application of Girsanov theorem,  $Q_{\mathbf{r}}^N(J)$  is absolutely continuous with respect to  $P_{\mathbf{r}} := \bigotimes_{i=1}^N P_{r_i}$ , and its density is given by the following equality:

$$(9) \quad \frac{dQ_{\mathbf{r}}^N(J)}{dP_{\mathbf{r}}}(\mathbf{x}) = \exp \left( \sum_{i=1}^N \int_0^T \left( \frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(X_{t-\tau(r_i, r_j)}^{j, N}) \right) dW_t(x^i, r_i) - \frac{1}{2} \int_0^T \left( \frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(X_{t-\tau(r_i, r_j)}^{j, N}) \right)^2 dt \right),$$

where

$$(10) \quad W_t(x, r) := \frac{x_t - x_0}{\lambda(r)} - \int_0^t \frac{f(x_s, r, s)}{\lambda(r)} ds.$$

Remark that, by (8),  $(W_t(\cdot, r))_t$  is a  $P_r$ -Brownian motion. Moreover, under  $P^{\otimes N}$  the Brownian motions  $(W_t(x^i, r_i))_t$  are independent.

Under  $P_{\mathbf{r}}$  neurons are independent but are not identically distributed as locations are heterogenous. We reduce this difficulty by averaging over locations: let  $P \in \mathcal{M}_1^+(\mathcal{C} \times D)$  as  $dP(x, r) := dP_r(x)d\pi(r)$ , the law of the pairs  $(X^{i, N}, r_i)$  when there is no interaction. This indeed properly defines a probability measure on  $\mathcal{M}_1^+(\mathcal{C} \times D)$  (see Appendix B). We also construct a symmetric law for the coupled network:

**Lemma 2.** *The map*

$$\mathcal{Q} : \begin{cases} D^N \rightarrow \mathcal{M}_1^+(\mathcal{C}^N) \\ \mathbf{r} \rightarrow Q_{\mathbf{r}}^N \end{cases}$$



where  $Q_{\mathbf{r}}^N := \mathcal{E}_J(Q_{\mathbf{r}}^N(J))$ , is continuous with respect to the weak topology. Moreover,

$$dQ^N(\mathbf{x}, \mathbf{r}) := dQ_{\mathbf{r}}^N(\mathbf{x}) d\pi^{\otimes N}(\mathbf{r})$$

defines a probability measure on  $\mathcal{M}_1^+((C \times D)^N)$ .

This result is proved in Appendix B.

*Remark 2.* • The probability measure  $Q^N$  averages the solutions on the different possible configurations  $(J, \mathbf{r})$ . Although being a relatively abstract object, it nevertheless provides relevant statistics as we make more explicit now. If  $A \subset (C \times D)^N$  is an event corresponding to e.g. a pathological behavior, then  $Q^N(A)$  corresponds to the proportion of configurations (“brains”) presenting this pathology. Conversely, as  $Q_{\mathbf{r}}^N(J)$  is the law of one particular individual with a given configuration  $(J, \mathbf{r})$ , then  $Q_{\mathbf{r}}^N(J)(A)$  provides the exact probability for him to suffer from  $A$ .

• Results under  $Q^N$  are called averaged, whereas those under  $Q_{\mathbf{r}}^N(J)$  are called quenched. Quenched results are much more involved to demonstrate than averaged ones. Several methods have been developed to access these results, particularly based on replica [5, 25, 4]. We do not address these questions in the present manuscript.

We are interested in the behavior of the double layer empirical measure:

$$(11) \quad \hat{\mu}_N = \frac{1}{N} \sum_{i=1}^N \delta_{(X^{i,N}, r_i)}.$$

Sanov’s theorem ensures that the empirical measure of the uncoupled network satisfies a full LDP. We will build upon this result and use an ad hoc version of Varadhan’s lemma to derive a weak LDP under  $Q^N$ . We will then characterize the possible minima of the associated good rate function, and prove by a fixed point argument that it admits a unique one, denoted  $Q$ , characterized as the non-Markovian solution of a MacKean-Vlasov SDE. Large deviations estimates will then ensure that the empirical measure converges exponentially fast toward this minimum. In detail, we show the following:

**Theorem 3.** *Under  $Q^N = \mathcal{E}(Q_{\mathbf{r}}^N(J))$  and for  $T < \frac{\lambda_*^2}{2\|\sigma\|_\infty^2}$ , the sequence of empirical measures  $(\hat{\mu}_N)_N$  satisfies a weak Large Deviations Principle of speed  $N$  and converges towards  $\delta_Q$  as  $N$  goes to infinity. Moreover, when delays are uniform in space, the upper-bound holds for any  $T > 0$ .*

*Remark 3.* Note that, for  $T < \frac{\lambda_*^2}{2\|\sigma\|_\infty^2}$ , a full large deviation principle can be demonstrated. Indeed, under the short-time hypothesis, we can readily prove exponential tightness of the averaged empirical measure, see e.g. as in [5]. In the present manuscript, we restrict our attention to the convergence of the empirical measure, thus the weak LDP suffices, and this allows combining our proofs with those in arbitrary time corresponding to the case of spatially uniform delays.

The quantitative estimates leading to this convergence result are summarized in the following two results:



**Theorem 4.** *There exists a good rate function  $H$  on  $\mathcal{M}_1^+(\mathcal{C} \times D)$  such that for any compact subset  $K$  of  $\mathcal{M}_1^+(\mathcal{C} \times D)$  and for  $T < \frac{\lambda_*^2}{2\|\sigma\|_\infty^2}$ ,*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in K) \leq -\inf_K H.$$

*Moreover, this upper-bound holds for any  $T > 0$  when delays are uniform in space.*

The convergence result also relies on the tightness of the sequence of empirical measures:

**Theorem 5.** *For any real number  $\varepsilon > 0$ , there exists a compact subset  $K_\varepsilon$  such that for any integer  $N$ ,*

$$Q^N(\hat{\mu}_N \notin K_\varepsilon) \leq \varepsilon.$$

These two results imply convergence of the empirical measure towards the set of minima of the rate function  $H$ . Their uniqueness and characterization is subject of the following theorem demonstrated in section 4:

**Theorem 6.** *The good rate function  $H$  achieves its minimal value at the unique probability measure  $Q \in \mathcal{M}_1^+(\mathcal{C} \times D)$  satisfying:*

$$Q \simeq P, \quad \frac{dQ}{dP}(x, r) = \mathcal{E} \left[ \exp \left\{ \frac{1}{\lambda(r)} \int_0^T G_t^Q(r) dW_t(x, r) - \frac{1}{2\lambda(r)^2} \int_0^T G_t^Q(r)^2 dt \right\} \right]$$

where  $(W_t)_{t \in [0, T]}$  is a  $P$ -brownian motion, and  $G^Q(r)$  is, under  $\mathcal{P}$ , a Gaussian process with mean and covariance

$$\begin{cases} \mathcal{E}[G_t^Q(r)] = \int_{\mathcal{C} \times D} J(r, r') S(x_{t-\tau(r, r')}) dQ(x, r') \\ \mathcal{E}[G_t^Q(r) G_s^Q(r)] = \int_{\mathcal{C} \times D} \sigma(r, r')^2 S(x_{t-\tau(r, r')}) S(x_{s-\tau(r, r')}) dQ(x, r). \end{cases}$$

For non-Gaussian synaptic weights, the LDP no more holds. Nevertheless, as for central limit theorems, the limit found in the Gaussian case is universal when the weights are sufficiently concentrated. Here, we will handle the case of weights having at most Gaussian tails:

$$(H_J) \quad \begin{cases} \exists a, D_0 > 0, \forall N \geq 1, \forall J_1 \in \{J_{ij}(N), i, j \in \llbracket 1, N \rrbracket\}, \\ \mathcal{E}_J \left( \exp \{a N J_1^2\} \right) \leq D_0. \end{cases}$$

We will show in section 5 that for times  $T < \lambda_*^2 a \wedge T^*$ , with  $T^* = \frac{\lambda_*^2}{2\|\sigma\|_\infty^2}$  in the general case and  $T^* = \infty$  for spatially homogeneous delays, the empirical measure converges exponentially fast towards the process described in Theorem 6. For sub-Gaussian synaptic weights (e.g., with bounded support), this convergence thus holds for any  $T < T^*$ . This indicates that the limit is universal to some degree. While condition  $(H_J)$  seems essential for having exponential speed of convergence, we expect that the universality of the Gaussian case goes beyond this case and may include synaptic weights having bounded polynomial moments (at least the two first moments). These extensions are not in the scope of the present paper, and our exponential convergence result covers all realistic cases arising in neuroscience where synaptic weights are bounded.

The convergence result of Theorem 3 actually also implies propagation of chaos, thanks to a very strong result due to Sznitman [38, Lemma 3.1]:

**Theorem 7.** *For any connectivity matrix satisfying hypothesis  $(H_J)$ , the system enjoys the propagation of chaos property. In other terms,  $Q^N$  is  $Q$ -chaotic, i.e. for any bounded continuous functions  $(\phi_1, \dots, \phi_m)$  and any neuron indexes  $(k_1, \dots, k_m)$ , we have:*

$$\lim_{N \rightarrow \infty} \int \prod_{j=1}^m \phi_j(x^{k_j}, r_{k_j}) dQ^N(x, r) = \prod_{j=1}^m \int \phi_j(x, r) dQ(x, r).$$

Throughout the paper, we will work in the whole space  $\mathcal{M}_1^+(\mathcal{C} \times D)$ . This is a very large set, containing in particular a number of measures that would not be satisfactory solutions for our biological problem. We will however show that the network equations converge to a limit  $Q$  that present all expected properties (section 4): it is absolutely continuous with respect to the law of an isolated neuron  $P$ , continuous in space (in a sense defined thereafter), defines a probability measure at every location, and presents  $\frac{1}{2}$ -Hölder trajectories in a  $L^2$ -sense.

We now proceed to the proof of our results.

### 3. LARGE DEVIATION PRINCIPLE

The aim of this section is to establish the weak large deviation principle for the network with Gaussian synaptic weights. It relies on three key points. First, we will characterize the good rate function; the intuition for constructing this functional comes from Varadhan's lemma. In our case it does not readily apply and we need to thoroughly demonstrate that the candidate is indeed a good rate function. Second, we will show an upper-bound result on compact sets. The spatially-extended framework will introduce new difficulties, spatially-dependent delays being particularly complex to handle. Nevertheless be able to cope with these while working with short times. Moreover, in the special case of uniform delays, strong estimates of the upper-bound are proven for general times. Third, the tightness of our collection of empirical measures will allow to conclude on a weak large-deviations principle.

**3.1. Construction of the good rate function.** Let us consider the interaction term of (2):

$$G_t^{N,i}(\mathbf{x}, \mathbf{r}) := \frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(x_{t-\tau(r_i, r_j)}^j).$$

As stated in section (1.2), it shall behave as a Gaussian process in the large  $N$  limit, with mean and covariance given by (3). With this in mind, we introduce, for  $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ , the two following functions defined respectively on  $[0, T]^2 \times D$  and  $[0, T] \times D$ :

$$\begin{cases} K_\mu(s, t, r) &:= \frac{1}{\lambda(r)^2} \int_{\mathcal{C} \times D} \sigma(r, r')^2 S(x_{t-\tau(r, r')}) S(x_{s-\tau(r, r')}) d\mu(x, r') \\ m_\mu(t, r) &:= \frac{1}{\lambda(r)} \int_{\mathcal{C} \times D} J(r, r') S(x_{t-\tau(r, r')}) d\mu(x, r'). \end{cases}$$

Here,  $\mu$  can be understood as the putative limit law of the couple  $(x^j, r_j)$  if it exists. Covariance and mean functions  $K_\mu$  and  $m_\mu$  are well defined as for every fixed  $r \in D$  the two maps

$$A_r : (x, r') \rightarrow \frac{1}{\lambda(r)} J(r, r') S(x_{t-\tau(r, r')}), \quad \tilde{A}_r(x, r') \rightarrow \frac{1}{\lambda(r)^2} \sigma(r, r')^2 S(x_{t-\tau(r, r')}) S(x_{s-\tau(r, r')})$$

are continuous for the classical product norm  $\|(x, r')\|_{\mathcal{C} \times D} := \sup_{t \in [-\tau, T]} |x(t)| + \|r'\|_{\mathbb{R}^d}$ .

Hence, they are Borel-measurable, and integrable with respect to every element of  $\mathcal{M}_1^+(\mathcal{C} \times D)$ . Remark that, since  $S$  takes value in  $[0, 1]$ , both functions are bounded:  $|K_\mu(s, t, r)| \leq \frac{\|\sigma\|_\infty^2}{\lambda_*^2}$  and  $|m_\mu(t, r)| \leq \frac{\|J\|_\infty}{\lambda_*}$ . Moreover, as  $\mu$  charges continuous functions,  $K_\mu$  and  $m_\mu$  are continuous maps by the dominated convergence theorem.

Clearly enough,  $K_\mu$  has a covariance structure. As a consequence, we are able to define for any measure  $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ , a family of independent stochastic processes  $(G(r))_{r \in D}$  as well as a probability measure  $\gamma_\mu$ , such that  $G(r)$  is a centered Gaussian process with covariance  $K_\mu(\cdot, \cdot, r)$  under  $\gamma_\mu$ . We will denote by  $\mathcal{E}_\mu$  the expectation under this measure.

*Remark 4.* As in [25], we could alternatively have defined a probability measure  $\gamma$ , and a family of Gaussian processes for this measure  $(G^\mu(r))_{\mu, r}$  independent for different locations, and with same means and covariances as above. This approach is equivalent to ours, but present the advantage of being very adapted to Fubini's theorem. For compactness of notation, we will mainly put the  $\mu$ -dependence in  $\gamma$ .

We recall a few general properties on the relative entropy that are often used throughout the paper. For  $p$  and  $q$  two probability measures on a Polish space  $E$  (see e.g. [20, Lemma 3.2.13]), we have the identity:

$$I(q|p) = \sup \left\{ \int_E \Phi dq - \log \int_E \exp \Phi dp ; \Phi \in \mathcal{C}_b(E) \right\},$$

which implies in particular that for any bounded measurable function  $\Phi$  on  $E$ ,

$$(12) \quad \int_{\mathcal{C}} \Phi dq \leq I(q|p) + \log \int_{\mathcal{C}} \exp \Phi dp.$$

If  $\Phi$  is a positive measurable function this inequality holds by monotone convergence, thus:

$$(13) \quad \int_{\mathcal{C}} \Phi dq \leq I(q|p) + \log \int_{\mathcal{C}} \exp \Phi dp.$$

We now state a key result to our analysis

**Lemma 8.**

$$\frac{dQ^N}{dP^{\otimes N}}(\mathbf{x}, \mathbf{r}) = \exp \left\{ N\Gamma(\hat{\mu}_N) \right\},$$

where

$$(14) \quad \Gamma(\mu) := \int_{\mathcal{C} \times D} \log \left( \int \exp \left\{ \int_0^T (G_t(\omega, r) + m_\mu(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(\omega, r) + m_\mu(t, r))^2 dt \right\} d\gamma_\mu(\omega) \right) d\mu(x, r),$$

and  $W(\cdot, r)$  is a Brownian motion under  $P_r$  defined in (10).

*Remark 5.* Since  $(x, r) \rightarrow (K_\mu(t, s, r), 0 \leq t, s \leq T), (m_\mu(t, r), 0 \leq t \leq T), (W_t(x, r), 0 \leq t \leq T)$  are continuous maps, the integral term in (14), thus  $\Gamma(\mu)$ , are well defined for any  $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ .

*Proof.* Let us go back to equation (9):

$$\frac{dQ_{\mathbf{r}}^N(J)}{dP_{\mathbf{r}}}(\mathbf{x}) = \exp \left( \sum_{i=1}^N \int_0^T G_t^{N,i}(\mathbf{x}, \mathbf{r}) dW_t(x^i, r_i) - \frac{1}{2} \int_0^T G_t^{N,i}(\mathbf{x}, \mathbf{r})^2 dt \right).$$

Averaging on  $J$  and applying Fubini theorem, we find that  $Q_{\mathbf{r}}^N \ll P_{\mathbf{r}}$ , with density

$$\frac{dQ_{\mathbf{r}}^N}{dP_{\mathbf{r}}}(\mathbf{x}) = \mathcal{E}_J \left[ \exp \left( \sum_{i=1}^N \int_0^T G_t^{N,i}(\mathbf{x}, \mathbf{r}) dW_t(x^i, r_i) - \frac{1}{2} \int_0^T G_t^{N,i}(\mathbf{x}, \mathbf{r})^2 dt \right) \right].$$

Moreover, equalities  $dQ^N(\mathbf{x}, \mathbf{r}) = dQ_{\mathbf{r}}^N(\mathbf{x}) d\pi^{\otimes N}(\mathbf{r})$ , and  $dP^{\otimes N}(\mathbf{x}, \mathbf{r}) = dP_{\mathbf{r}}(\mathbf{x}) d\pi^{\otimes N}(\mathbf{r})$  give

$$\frac{dQ^N}{dP^{\otimes N}}(\mathbf{x}, \mathbf{r}) = \prod_{i=1}^N \mathcal{E}_J \left[ \exp \left( \int_0^T G_t^{N,i}(\mathbf{x}, \mathbf{r}) dW_t(x^i, r_i) - \frac{1}{2} \int_0^T G_t^{N,i}(\mathbf{x}, \mathbf{r})^2 dt \right) \right],$$

where we used the independence of the synaptic weights  $J_{ij}$ . Note that here  $\mathbf{x}$  are coordinates, thus independent of the  $J_{ij}$ . Moreover  $\left\{ \frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(x_{t-\tau(r_i, r_j)}^j), 0 \leq t \leq T \right\}$  is, under  $\mathcal{P}_J$ , a Gaussian process with covariance  $K_{\hat{\mu}_N}(t, s, r_i)$ , and mean  $m_{\hat{\mu}_N}(t, r_i)$ , we can replace it by  $G_t(r_i) + m_{\hat{\mu}_N}(t, r_i)$  taken under  $\gamma_{\hat{\mu}_N}$ :

$$\begin{aligned} \frac{dQ^N}{dP^{\otimes N}}(\mathbf{x}, \mathbf{r}) &= \exp \left\{ \sum_{i=1}^N \log \left( \int \exp \left\{ \int_0^T (G_t(r_i) + m_{\hat{\mu}_N}(t, r_i)) dW_t(x^i, r_i) \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{1}{2} \int_0^T (G_t(r_i) + m_{\hat{\mu}_N}(t, r_i))^2 dt \right\} d\gamma_{\hat{\mu}_N} \right) \right\} \\ &= \exp \left\{ N \int \log \left( \int \exp \left\{ \int_0^T (G_t(r) + m_{\hat{\mu}_N}(t, r)) dW_t(x, r) \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{1}{2} \int_0^T (G_t(r) + m_{\hat{\mu}_N}(t, r))^2 dt \right\} d\gamma_{\hat{\mu}_N} \right) d\hat{\mu}_N(x, r) \right\}, \end{aligned}$$

which concludes the proof.  $\square$

The function  $\Gamma$  has the following properties:

**Proposition 9.** (1)  $\Gamma \leq I(\cdot|P)$ . In particular,  $\Gamma$  is finite whenever  $I(\cdot|P)$  is.  
(2)  $\exists \iota < 1, e > 0$ , such that  $\Gamma \leq \iota I(\cdot|P) + e$ .

*Proof.* (1):

Let  $F_{\mu}$  denote the integrand in the formulation of  $\Gamma$  (14):

$$F_{\mu}(x, r) := \log \left\{ \int \exp \left\{ \int_0^T (G_t(r) + m_{\mu}(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_{\mu}(t, r))^2 r dt \right\} d\gamma_{\mu} \right\},$$

and  $F_{\mu, M}$  the following regularization of  $F_{\mu}$ :

$$F_{\mu, M}(x, r) := \log \left\{ \int M \wedge \exp \left\{ \int_0^T (G_t(r) + m_{\mu}(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_{\mu}(t, r))^2 r dt \right\} d\gamma_{\mu} \right\}.$$

The latter functional is positive bounded and measurable, thus inequality (13) holds. By the monotone convergence theorem, one obtains for every  $\alpha \geq 1$ ,

$$\begin{aligned}
\alpha \int_{\mathcal{C} \times D} F_\mu(x, r) d\mu(x, r) &\leq I(\mu|P) + \log \left\{ \int \exp \alpha F_\mu(x, r) dP \right\} \\
&\stackrel{\text{Jensen}}{\leq} I(\mu|P) + \log \left\{ \int \int \exp \left\{ \alpha \int_0^T (G_t(r) + m_\mu(t, r)) dW_t(x, r) \right. \right. \\
&\quad \left. \left. - \frac{\alpha}{2} \int_0^T (G_t(r) + m_\mu(t, r))^2 dt \right\} d\gamma_\mu dP(x, r) \right\} \\
&\stackrel{\text{Fubini}}{\leq} I(\mu|P) + \log \left\{ \int \int_D \int_{\mathcal{C}} \exp \left\{ \alpha \int_0^T (G_t(r) + m_\mu(t, r)) dW_t(x, r) \right. \right. \\
&\quad \left. \left. - \frac{\alpha}{2} \int_0^T (G_t(r) + m_\mu(t, r))^2 dt \right\} dP_r(x) d\pi(r) d\gamma_\mu \right\}.
\end{aligned}$$

Then  $W(\cdot, r)$  being a  $P_r$ -Brownian motion, the martingale property yields

$$\alpha \int_{\mathcal{C} \times D} F_\mu(x, r) d\mu(x, r) \leq I(\mu|P) + \log \left\{ \int \int \exp \left\{ \frac{\alpha^2 - \alpha}{2} \int_0^T (G_t(r) + m_\mu(t, r))^2 dt \right\} d\gamma_\mu d\pi(r) \right\},$$

from which we conclude by letting  $\alpha = 1$ .

**(2):**

Basic Gaussian calculus gives

(15)

$$\mathbb{E} \left[ \exp \left\{ \frac{1}{2} \mathcal{N}(m, v)^2 \right\} \right] = \frac{1}{\sqrt{1-v}} \exp \left\{ \frac{m^2}{2(1-v)} \right\} = \exp \left\{ \frac{1}{2} \left( \frac{m^2}{1-v} - \log(1-v) \right) \right\}$$

as soon as  $v < 1$ . Jensen inequality and Fubini theorem yield

$$\int \exp \left\{ \frac{(\alpha^2 - \alpha)T}{2} \int_0^T (G_t(r) + m_\mu(t, r))^2 \frac{dt}{T} \right\} d\gamma_\mu \leq \int_0^T \int \exp \left\{ \frac{(\alpha^2 - \alpha)T}{2} (G_t(r) + m_\mu(t, r))^2 \right\} d\gamma_\mu \frac{dt}{T}.$$

$$\text{As } \sqrt{(\alpha^2 - \alpha)T} (G_t(r) + m_\mu(t, r)) \sim \mathcal{N} \left( \sqrt{(\alpha^2 - \alpha)T} m_\mu(t, r), (\alpha^2 - \alpha)TK_\mu(t, t, r) \right)$$

under  $\gamma_\mu$  then, for  $(\alpha - 1)$  small enough, exists a constant  $C_T$  uniform in space such that

$$\begin{aligned}
\int \exp \left\{ \frac{(\alpha^2 - \alpha)}{2} \int_0^T (G_t(r) + m_\mu(t, r))^2 dt \right\} d\gamma_\mu &\leq \exp \left\{ (\alpha - 1)C_T + \underbrace{o(\alpha - 1)}_{\text{uniform in } r} \right\} \\
&\leq \exp \left\{ (\alpha - 1)C_T \right\}.
\end{aligned}$$

This eventually yields

$$\Gamma(\mu) \leq \iota I(\mu|P) + e,$$

with  $\iota := \frac{1}{\alpha}$ , and  $e := (\alpha - 1)C_T$ .

□

As  $\mathcal{C} \times D$  and  $\mathcal{M}_1^+(\mathcal{C} \times D)$  are Polish spaces, and as the  $(X^{i,N}, r_i)$  are independent identically distributed random variables under  $P^{\otimes N}$ , Sanov's theorem ensures that the empirical measure satisfies, under this measure, a LDP with good rate function  $I(\cdot|P)$ . Furthermore, if  $\Gamma$  was bounded and continuous, Varadhan's lemma would, as a consequence of Lemma (8), entail a full LDP under  $Q^N$ , with good rate function given by

$$H(\mu) := \begin{cases} I(\mu|P) - \Gamma(\mu) & \text{if } I(\mu|P) < \infty, \\ \infty & \text{otherwise .} \end{cases}$$

At this point, it would be easy to conclude provided that  $\Gamma$  presented a few regularity properties. Unfortunately, Varadhan's lemma assumptions fail here. Obtaining a weak LDP as well as the convergence of the empirical measure requires to come back to the basics of large deviations theory.

Observe that  $\Gamma$  is a nonlinear function of  $\mu$ , involving in particular an exponential term depending on the mean and covariance structure of the Gaussian process. In order to handle terms of this type, a key technique proposed by Ben Arous and Guionnet is to linearize this map by considering the terms in the exponential as depending on an additional variable  $\nu$  [5, 25]. In our case, as family of linearizations are given by the maps:

$$\Gamma_\nu(\mu) := \int_{\mathcal{C} \times D} \log \left( \int_0^T \exp \left\{ \int_0^T (G_t(\omega, r) + m_\nu(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(\omega, r) + m_\nu(t, r))^2 dt \right\} d\gamma_\nu(\omega) \right) d\mu(x, r),$$

where  $\mu, \nu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ . Moreover, we note that  $\Gamma_\nu(\hat{\mu}_N) = \frac{1}{N} \sum_{i=1}^N \Gamma_\nu(\delta_{(x^i, r_i)})$ , and introduce the notation:

$$\exp \{ N \Gamma_\nu(\hat{\mu}_N) \} dP^{\otimes N}(\mathbf{x}, \mathbf{r}) = \left( \exp \{ \Gamma_\nu(\delta_{(x, r)}) \} dP(x, r) \right)^{\otimes N} =: dQ_\nu(x, r)^{\otimes N}.$$

This equality highlights the fact that, applying again Sanov theorem, the empirical measure satisfies a full LDP under  $(Q_\nu)^{\otimes N}$ , with good rate function  $I(\cdot|Q_\nu)$ . On the other hand, Vardhan's lemma suggests that  $\hat{\mu}_N$  satisfies, under the same measure, a LDP with rate function

$$H_\nu : \begin{cases} \mathcal{M}_1^+(\mathcal{C} \times D) & \rightarrow \mathbb{R}^+ \\ \mu & \rightarrow \begin{cases} I(\mu|P) - \Gamma_\nu(\mu) & \text{if } I(\mu|P) < \infty, \\ \infty & \text{otherwise .} \end{cases} \end{cases}$$

This is, for now, only a supposition, as its original counterpart,  $\Gamma_\nu$ , is not bounded from above nor continuous. Still, assuming the result is true, uniqueness of the good rate function would imply that  $H_\nu$  equals  $I(\cdot|Q_\nu)$ . Before justifying the definition of  $Q_\nu$ , and proceeding to the rigorous demonstration of the latter equality, we introduce a useful decomposition of  $\Gamma_\nu$ .

As demonstrated in [13], we have:

**Proposition 10.** *For every  $\nu, \mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ ,  $\Gamma_\nu$  admits the following decomposition:*

$$\Gamma_\nu(\mu) = \Gamma_{1,\nu}(\mu) + \Gamma_{2,\nu}(\mu).$$

where

$$\begin{aligned}\Gamma_{1,\nu}(\mu) &:= \int_{\mathcal{C} \times D} \left\{ \log \left( \int \exp \left( -\frac{1}{2} \int_0^T G_t(r)^2 dt \right) d\gamma_\nu \right) - \frac{1}{2} \int_0^T m_\nu(t, r)^2 dt \right\} d\mu(x, r), \\ \Gamma_{2,\nu}(\mu) &:= \frac{1}{2} \int \int \left( \int G_t(r) (dW_t(x, r) - m_\nu(t, r) dt) \right)^2 d\gamma_{\tilde{K}_{\nu,r}^T} d\mu(x, r) \\ &\quad + \int \int m_\nu(t, r) dW_t(x, r) d\mu(x, r),\end{aligned}$$

and

$$d\gamma_{\tilde{K}_{\nu,r}^T} := \frac{\exp \left\{ -\frac{1}{2} \int_0^T (G_t(\omega, r))^2 dt \right\}}{\int \exp \left\{ -\frac{1}{2} \int_0^T (G_t(\omega, r))^2 dt \right\} d\gamma_\nu} d\gamma_\nu.$$

*Remark 6.* • As proven in [5, Appendix A],  $\gamma_{\tilde{K}_{\nu,r}^T}$  is a probability measure on  $\Omega$  under which  $G(r)$  is a centered Gaussian process with covariance

$$\tilde{K}_{\nu,r}^t(s, u) := \left( \int \frac{\exp \left\{ -\frac{1}{2} \int_0^t (G_u(\omega, r))^2 du \right\} G_u(\omega, r) G_s(\omega, r)}{\int \exp \left\{ -\frac{1}{2} \int_0^t (G_u(\omega, r))^2 du \right\} d\gamma_\nu} d\gamma_\nu \right).$$

- This decomposition has the interest of splitting the difficulties: while the first term will be relatively easy to handle, the local martingale term will require finer estimates based on Gaussian calculus and a number of tools from stochastic calculus theory.

The decomposition is particularly useful to prove the following:

**Theorem 11.**  $Q_\nu$  is a well defined probability measure on  $\mathcal{M}_1^+(\mathcal{C} \times D)$ , and  $H_\nu(\mu) = I(\mu|Q_\nu)$ . In particular  $H_\nu$  is a good rate function.

*Proof.* By definition of  $\Gamma_\nu$ , on has

$$\begin{aligned}dQ_\nu(x, r) &:= \exp \{ \Gamma_\nu(\delta_x \otimes \delta_r) \} dP(x, r) \\ &= \int \exp \left( \int_0^T (G_t(r) + m_\nu(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_\nu(t, r))^2 dt \right) d\gamma_\nu dP(x, r)\end{aligned}$$

We deduce by the martingale property of its density that it is in fact a probability measure. We will first prove that  $I(Q_\nu|P)$  is finite. To this purpose, we introduce, for every fixed  $r \in D$ , the probability  $Q_{\nu,r}$  as follow:

$$dQ_{\nu,r}(x) := \exp \Gamma_\nu(\delta_x \otimes \delta_r) dP_r(x).$$

By martingale property, it is well defined. As in [5, Lemma 5.15], we can show that there exists an adapted process  $(H_\nu(t, x, r))$  such that:

$$\frac{dQ_{\nu,r}}{dP_r}(x) = \exp \left\{ \int_0^T H_\nu(t, x, r) dW_t(x, r) - \frac{1}{2} \int_0^T H_\nu^2(t, x, r) dt \right\},$$

where

$$H_\nu(t, x, r) = \int_0^t \tilde{K}_{\nu,r}^t(t, s) (dW_s(x, r) - m_\nu(s, r) ds) + m_\nu(t, r).$$



Girsanov theorem ensures the existence of a  $Q_{\nu,r}$ -Brownian motion  $(B_t(\cdot, r))$  such that  $W_t(x, r) = B_t(x, r) + \int_0^t H_\nu(s, x, r)ds$ , yielding

$$\begin{aligned}
I(Q_{\nu,r}|P_r) &= \int_{\mathcal{C}} \left\{ \int_0^T H_\nu(t, x, r) dW_t(x, r) - \frac{1}{2} \int_0^T H_\nu^2(t, x, r) dt \right\} dQ_{\nu,r}(x) \\
&= \int_{\mathcal{C}} \left\{ \int_0^T H_\nu(t, x, r) dB_t(x, r) + \frac{1}{2} \int_0^T H_\nu^2(t, x, r) dt \right\} dQ_{\nu,r}(x) = \frac{1}{2} \int_{\mathcal{C}} \left\{ \int_0^T H_\nu^2(t, x, r) dt \right\} dQ_{\nu,r}(x) \\
&\leq \int_{\mathcal{C}} \int_0^T \left\{ \left( \int_0^t \tilde{K}_{\nu,r}^t(t, s) (dW_s(x, r) - m_\nu(s, r)ds) \right)^2 + m_\nu(t, r)^2 \right\} dt dQ_{\nu,r}(x) \\
&\leq \int_0^T \underbrace{\left\{ \int_{\mathcal{C}} \left( \int_0^t \tilde{K}_{\nu,r}^t(t, s) (dW_s(x, r) - m_\nu(s, r)ds) \right)^2 dQ_{\nu,r}(x) \right\}}_{\varphi_\nu(t, r)} dt + \frac{T \|\bar{J}\|_\infty^2}{\lambda_*^2}.
\end{aligned}$$

As in [13, Theorem 2], Gronwall's lemma yields

$$\sup_{t \leq T} \varphi_\nu(t, r) \leq 2 \frac{\|\sigma\|_\infty^4 T}{\lambda_*^4} \exp \left\{ 2 \frac{\|\sigma\|_\infty^4 T}{\lambda_*^4} \right\},$$

which implies:

$$I(Q_{\nu,r}|P_r) \leq 2 \frac{\|\sigma\|_\infty^4 T^2}{\lambda_*^4} e^{2 \frac{\|\sigma\|_\infty^4 T}{\lambda_*^4}} + \frac{\|\bar{J}\|_\infty^2 T}{\lambda_*^2}.$$

Moreover,  $I(Q_{\nu,r}|P_r)$  is positive and  $\pi$ -measurable. We thus integrate on  $D$ :

$$I(Q_\nu|P) = \int_{\mathcal{C} \times D} I(Q_{\nu,r}|P_r) d\pi(r) \leq 2 \frac{\|\sigma\|_\infty^4 T^2}{\lambda_*^4} e^{2 \frac{\|\sigma\|_\infty^4 T}{\lambda_*^4}} + \frac{\|\bar{J}\|_\infty^2 T}{\lambda_*^2} < \infty.$$

Let us now prove the announced identity. We define, for every  $\mu \in \mathcal{M}_1^+(\mathcal{C})$  and  $r \in D$ ,

$$\Gamma_{\nu,r}(\mu) := \int_{\mathcal{C}} \log \left( \int \exp \left\{ \int_0^T (G_t(r) + m_\nu(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_\nu(t, r))^2 dt \right\} d\gamma_\nu \right) d\mu(x),$$

and

$$H_{\nu,r}(\mu) := I(\mu|P_r) - \Gamma_{\nu,r}(\mu).$$

For fixed  $r$ , we can prove as in Lemma 5.(v) of [13] that

$$(16) \quad H_{\nu,r}(\cdot) = I(\cdot|Q_{\nu,r}).$$

Let now  $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ . If  $\mu \not\ll P$ , then  $H_\nu(\mu) = \infty$ . This also implies that  $\mu \not\ll Q_\nu$  as  $Q_\nu \ll P$ , so that  $I(\mu|Q_\nu) = +\infty$ . Suppose now that  $\mu \ll P$ . This implies that  $\mu$  has a measurable density  $\rho_\mu$  with respect to  $\mathcal{B}(\mathcal{C} \times D)$ :

$$d\mu(x, r) = \rho_\mu(x, r) dP(x, r) = \rho_\mu(x, r) dP_r(x) d\pi(r).$$

Hence, for  $r \in D$  such that  $c_\mu(r) := \int_D \rho_\mu(x, r) dP_r(x) \neq 0$ , we can properly define  $\mu_r \in \mathcal{M}_1^+(\mathcal{C})$  by  $d\mu_r(x) := \frac{\rho_\mu(x, r)}{c_\mu(r)} dP_r(x)$ . Of course  $\mu_r \ll P_r$ , and

$$(17) \quad d\mu(x, r) = d\mu_r(x) c_\mu(r) d\pi(r).$$

Remark that  $c_\mu$  is a measurable function in space, and that the set  $\{r \in D, c_\mu(r) = 0\}$  will not impact the value of the integral of interest. Then, for such a  $\mu$

$$\Gamma_\nu(\mu) = \int_D \Gamma_{\nu,r}(\mu_r) c_\mu(r) d\pi(r), \quad I(\mu|P) = \int_D I(\mu_r|P_r) c_\mu(r) d\pi(r) + \int_D \log(c_\mu(r)) c_\mu(r) d\pi(r),$$

with both sides possibly infinite. Combining these two equalities as well as (16), we conclude that

$$\begin{aligned} H_\nu(\mu) &= \int_D H_{\nu,r}(\mu_r) c_\mu(r) d\pi(r) + \int_D \log(c_\mu(r)) c_\mu(r) d\pi(r) \\ &= \int_D I(\mu_r|Q_{\nu,r}) c_\mu(r) d\pi(r) + \int_D \log(c_\mu(r)) c_\mu(r) d\pi(r) = I(\mu|Q_\nu). \end{aligned}$$

□

We have thus proved that  $H_\nu$  is a good rate function, and would like to extend this property to  $H$ :  $H_\nu$  is seen in our proof as an intermediate tool, equal to  $H$  when  $I(\mu|P) = \infty$ , but differing of  $\Gamma - \Gamma_\nu$  otherwise. We control this difference below in Lemma 12.

Let us introduce two preliminary objects that will appear in the obtained upper-bound. First, define the Vaserstein-like distance on  $\mathcal{M}_1^+(\mathcal{C} \times D)$ , compatible with the weak topology:

$$d_T(\mu, \nu) := \inf_{\xi} \left\{ \int \left( \sup_{-\tau \leq t \leq T} |x_t - y_t|^2 + |r - r'| \right) d\xi((x, r), (y, r')) \right\}^{\frac{1}{2}}$$

the infimum being taken on the laws  $\xi$  with marginals  $\mu$  and  $\nu$ . Second, we define the following  $L^2$  quantifier of Hölder  $\frac{1}{2}$  continuity:

$$\chi : \begin{cases} \mathcal{M}_1^+(\mathcal{C} \times D) \rightarrow \mathbb{R}^+ \cup \{+\infty\} \\ \mu \rightarrow \sup_{s, t \in [-\tau, T], s \neq t} \int_{\mathcal{C} \times D} \frac{|x_t - x_s|^2}{|t - s|} d\mu(x, r). \end{cases}$$

We have now introduced all elements to state the following technical lemma concluding on the fact that  $H$  is a good rate function.

**Lemma 12.** *Let  $\mu, \nu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ , then:*

- (1) *there exists real constants  $b, c > 0$ , such that  $\chi(\mu) \leq bI(\mu|P) + c$ . In particular,  $\chi$  is finite whenever  $I(\cdot|P)$  is.*
- (2) *there exists a positive constant  $C_T$  such that:*
  - (a)  $|\Gamma_{1,\nu}(\mu) - \Gamma_1(\mu)| \leq C_T \sqrt{1 + \chi(\mu)} d_T(\mu, \nu)$ .
  - (b)  $|\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)| \leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu|P)) d_T(\mu, \nu)$ .
- (3)  *$H$  is a good rate function.*

*Proof.* The main techniques were introduced in [5, Lemma 3.3-3.4] and used in a neuroscience setting in [13, Lemma.5], but spatiality induces new difficulties. In particular, the regularity in space and the map  $\chi$  are absent in these previous contributions. We focus on the main challenges of the spatial nature of the equation, and otherwise refer to these two references.

**Proof of Lemma 12.(1).**

Using (13), one finds that for every  $\varepsilon > 0, s < t \in [-\tau, T]$ :

$$(18) \quad \int_{\mathcal{C} \times D} \varepsilon \frac{|x_t - x_s|^2}{|t - s|} d\mu(x, r) \leq I(\mu|P) + \log \left( \int_D \int_{\mathcal{C}} \exp \left\{ \varepsilon \frac{|x_t - x_s|^2}{|t - s|} \right\} dP_r(x) d\pi(r) \right).$$

Set  $s \in [-\tau, T[$ . Under  $P_r$ ,

$$\begin{aligned} \underbrace{\int_s^t |f(r, u, x_u)| du}_{g_s(t)} &\leq \int_s^t |f(r, s, x_s)| du + \int_s^t |f(r, u, x_u) - f(r, s, x_s)| du \\ &\leq |f(r, s, x_s)|(t-s) + K_f \int_s^t \{(u-s) + |x_u - x_s|\} du \\ &\leq |f(r, s, x_s)|(t-s) + K_f \frac{(t-s)^2}{2} + K_f \int_s^t g_s(u) du + K_f \lambda^* \int_s^t |W_u(x, r) - W_s(x, r)| du. \end{aligned}$$

Gronwall's lemma yields

$$(19) \quad \int_s^t |f(r, u, x_u)| du \leq \left( |f(r, s, x_s)|(t-s) + K_f \frac{(t-s)^2}{2} + K_f \lambda^* \int_s^t |W_u(x, r) - W_s(x, r)| du \right) \exp \{K_f(t-s)\}.$$

Now,

$$\begin{aligned} \varepsilon \frac{|x_t - x_s|^2}{|t-s|} &\leq 2\varepsilon \frac{|\int_s^t f(r, u, x_u) du|^2}{|t-s|} + 2\varepsilon \lambda(r)^2 \left( \frac{W_t(x, r) - W_s(x, r)}{\sqrt{t-s}} \right)^2 \\ &\stackrel{C.S.}{\leq} 4\varepsilon \exp\{2K_f(t-s)\} \left\{ (t-s) \left( |f(r, s, x_s)| + K_f \frac{(t-s)}{2} \right)^2 \right. \\ &\quad \left. + (K_f \lambda^*)^2 \frac{1}{(t-s)} \left( \int_s^t |W_u(x, r) - W_s(x, r)| du \right)^2 \right\} + 2\varepsilon \lambda(r)^2 \left( \frac{W_t(x, r) - W_s(x, r)}{\sqrt{t-s}} \right)^2 \\ &\stackrel{Jensen}{\leq} \varepsilon C_T \left\{ \left( |f(r, s, x_s)| + 1 \right)^2 + \int_s^t |W_u(x, r) - W_s(x, r)|^2 du + \left( \frac{W_t(x, r) - W_s(x, r)}{\sqrt{t-s}} \right)^2 \right\} \end{aligned}$$

so that

$$\begin{aligned} \int_{\mathcal{C}} \exp \left\{ \varepsilon \frac{|x_t - x_s|^2}{|t-s|} \right\} dP_r(x) &\stackrel{\text{Hölder}}{\leq} \left\{ \exp \left\{ \varepsilon C_T \left( |f(r, s, x_s)| + 1 \right)^2 \right\} dP_r(x) \right\}^{\frac{1}{3}} \\ &\times \left\{ \int_{\mathcal{C}} \exp \left\{ \varepsilon C_T \int_s^t (W_u(x, r) - W_s(x, r))^2 du \right\} dP_r(x) \right\}^{\frac{1}{3}} \left\{ \int_{\mathcal{C}} \exp \left\{ \varepsilon C_T \left( \frac{W_t(x, r) - W_s(x, r)}{\sqrt{t-s}} \right)^2 \right\} dP_r(x) \right\}^{\frac{1}{3}}. \end{aligned}$$

For  $\varepsilon$  small enough, the first integrand is finite (see Appendix, Lemma 25). For the other two Gaussian terms, Jensen's inequality on one hand, and (15) on the other hand, imply that  $\varepsilon$  only needs to be inferior to  $\frac{1}{2C_T(t-s)}$ , and  $\frac{1}{2C_T}$  respectively, in order for them to be finite. Thus, there exists a constant  $C$ , uniform in space such that

$$\int_{\mathcal{C}} \exp \left\{ \varepsilon \frac{|x_t - x_s|^2}{|t-s|} \right\} dP_r(x) \leq C.$$

Using (18), we conclude that  $\chi(\mu) \leq \frac{1}{\varepsilon} I(\mu|P) + \frac{\log(C)}{\varepsilon}$ .

**Proof of Lemma 12.(2.a).**

In order to get the reader used to spatiality, we give the full proof this point, while the others shall be omitted when no new difficulties emerge.

We define

$$\Gamma_1(\mu, r) := \left\{ \log \left( \int \exp \left( -\frac{1}{2} \int_0^T G_t(r)^2 dt \right) d\gamma_\mu \right) - \frac{1}{2} \int_0^T m_\mu(t, r)^2 dt \right\},$$

so that

$$\Gamma_{1,\nu}(\mu) - \Gamma_1(\mu) = \int_{\mathcal{C} \times D} (\Gamma_1(\nu, r) - \Gamma_1(\mu, r)) d\mu(x, r).$$

Jensen inequality gives a lowerbound of the following quantity:

$$\begin{aligned} \Gamma_1(\mu, r) + \frac{1}{2} \int m_\mu^2(t, r) dt &= \log \int \exp \left\{ -\frac{1}{2} \int_0^T G_t^2(r) dt \right\} d\gamma_\mu \\ &\geq -\frac{1}{2} \int \left\{ \int_0^T G_t^2(r) dt \right\} d\gamma_\mu \geq -\frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2}. \end{aligned}$$

Remark that it does not depend upon space, so that, starting with estimates at fixed position  $r \in D$ , we will be able to use most lower- and upper-bounds in the non-spatial case [13, Lemma 5]. Following painstakingly these proofs, we obtain for a fixed location  $r \in D$ ,

$$\begin{aligned} \left| \Gamma_1(\mu, r) - \Gamma_1(\nu, r) + \frac{1}{2} \int (m_\mu^2 - m_\nu^2)(t, r) dt \right| &= \left| \log \left( 1 + \frac{\int \exp \left( -\frac{1}{2} \int_0^T G_t^2(r) dt \right) d(\gamma_\mu - \gamma_\nu)}{\int \exp \left( -\frac{1}{2} \int_0^T G_t^2(r) dt \right) d\gamma_\nu} \right) \right| \\ &\leq \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\} \left| \int \exp \left( -\frac{1}{2} \int_0^T G_t^2(r) dt \right) d(\gamma_\mu - \gamma_\nu) \right|, \end{aligned}$$

Let  $\xi$  be a probability measure on  $(\mathcal{C} \times D) \times (\mathcal{C} \times D)$  with marginals  $\mu$  and  $\nu$ , and let  $(G(r), G'(r))_{r \in D}$  be, under  $\gamma_\xi$ , a family of independent bidimensional centered Gaussian processes with covariance  $K_\xi(s, t, r)$  given by:

$$(20) \quad \frac{1}{\lambda(r)^2} \int \begin{pmatrix} \sigma_{rr'}^2 S(x_{s-\tau_{rr'}}) S(x_{t-\tau_{rr'}}) & \sigma_{rr'} \sigma_{r\tilde{r}'} S(x_{s-\tau_{rr'}}) S(y_{t-\tau_{r\tilde{r}'}}) \\ \sigma_{rr'} \sigma_{r\tilde{r}'} S(y_{s-\tau_{r\tilde{r}'}}) S(x_{t-\tau_{rr'}}) & \sigma_{r\tilde{r}'}^2 S(y_{s-\tau_{r\tilde{r}'}}) S(y_{t-\tau_{r\tilde{r}'}}) \end{pmatrix} d\xi((x, r'), (y, \tilde{r}')).$$

with the short-hand notations  $\sigma_{rr'} := \sigma(r, r')$ ,  $\tau_{rr'} = \tau(r, r')$ .

Then,

$$\begin{aligned} \left| \int \exp \left( -\frac{1}{2} \int_0^T G_t^2(r) dt \right) d(\gamma_\mu - \gamma_\nu) \right| &= \left| \int \left\{ \exp \left( -\frac{1}{2} \int_0^T G_t^2(r) dt \right) - \exp \left( -\frac{1}{2} \int_0^T G_t'^2(r) dt \right) \right\} d\gamma_\xi \right| \\ &\leq \frac{1}{2} \int \int_0^T |G_t^2(r) - G_t'^2(r)| dt d\gamma_\xi \\ &\stackrel{C.S.}{\leq} \frac{1}{2} \prod_{\varepsilon=\pm 1} \left( \int \int_0^T (G_t(r) + \varepsilon G_t'(r))^2 dt d\gamma_\xi \right)^{\frac{1}{2}}. \end{aligned}$$

Then, using the covariance of  $(G(r), G'(r))$  under  $\gamma_\xi$ , one finds:

(21)

$$\begin{aligned} \left| \Gamma_1(\mu, r) - \Gamma_1(\nu, r) + \frac{1}{2} \int (m_\mu^2 - m_\nu^2)(t, r) dt \right| &\leq \frac{1}{2} \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\} \left( \frac{4\|\sigma\|_\infty^2 T}{\lambda_*^2} \right)^{\frac{1}{2}} \\ &\times \left\{ \frac{1}{\lambda_*^2} \int \int_0^T \left( \sigma(r, r') S(x_{t-\tau(r, r')}) - \sigma(r, \tilde{r}') S(y_{t-\tau(r, \tilde{r}')}) \right)^2 dt d\xi((x, r'), (y, \tilde{r}')) \right\}^{\frac{1}{2}}. \end{aligned}$$

Here appears the first difficulty associated with spatial dimension. Since variance and delays depend on locations, we are not able to rely only on the Lipschitz continuity of  $S$  to conclude: regularity in space is necessary, involving in particular  $\chi(\mu)$ . In detail, splitting the integrand of the righthand side of (21), we find:

$$\begin{aligned} & \left( \sigma_{rr'} S(x_{t-\tau_{rr'}}) - \sigma_{r\tilde{r}'} S(y_{t-\tau_{r\tilde{r}'}}) \right)^2 \leq 2 \left\{ (\sigma_{rr'} - \sigma_{r\tilde{r}'})^2 S(x_{t-\tau_{rr'}})^2 + \sigma_{r\tilde{r}'}^2 (S(x_{t-\tau_{rr'}}) - S(y_{t-\tau_{r\tilde{r}'}}))^2 \right\} \\ & \leq 2K_\sigma^2 |r' - \tilde{r}'|^2 + 4\|\sigma\|_\infty^2 \left\{ (S(x_{t-\tau_{rr'}}) - S(x_{t-\tau_{r\tilde{r}'}}))^2 + (S(x_{t-\tau_{r\tilde{r}'}}) - S(y_{t-\tau_{r\tilde{r}'}}))^2 \right\} \\ & \leq 2K_\sigma^2 d_D |r' - \tilde{r}'| + 4\|\sigma\|_\infty^2 \left( K_S |x_{t-\tau_{rr'}} - x_{t-\tau_{r\tilde{r}'}}| \right)^2 + 4\|\sigma\|_\infty^2 K_S^2 \sup_{s \in [-\tau, T]} |x_s - y_s|^2, \end{aligned}$$

where  $d_D := \text{diam}(D)$ . As  $\int_{\mathcal{C} \times D} (x_t - x_s)^2 d\mu(x, r) \leq \chi(\mu) |t - s|$ , and as  $\tau$  is Lipschitz continuous in its second variable, one finds:

$$(22) \quad \left| \Gamma_1(\mu, r) - \Gamma_1(\nu, r) + \frac{1}{2} \int (m_\mu^2 - m_\nu^2)(t, r) dt \right| \leq C_T \sqrt{1 + \chi(\mu)} d_T(\mu, \nu).$$

Moreover, one has:

$$\begin{aligned} \left| \int (m_\mu^2 - m_\nu^2)(t, r) dt \right| &= \int \left| (m_\mu - m_\nu)(m_\mu + m_\nu) \right| (t, r) dt \\ &\leq 2 \frac{\|\bar{J}\|_\infty}{\lambda_*} \int |(m_\mu - m_\nu)(t, r)| dt, \end{aligned}$$

and

$$\begin{aligned} & \int_0^T |(m_\mu - m_\nu)(t, r)| dt = \int_0^T \left| \frac{1}{\lambda(r)} \int J(r, r') S(x_{t-\tau(r, r')}) d(\mu - \nu)(x, r') \right| dt \\ & \leq \frac{1}{\lambda_*} \int_0^T \int_0^T \left| J(r, r') S(x_{t-\tau(r, r')}) - J(r, \tilde{r}') S(y_{t-\tau(r, \tilde{r}')} ) \right| dt d\xi((x, r'), (y, \tilde{r}')) \\ & \leq \frac{1}{\lambda_*} \left\{ K_J T \int |\tilde{r}' - r'| d\xi((x, r'), (y, \tilde{r}')) + \|\bar{J}\|_\infty \int \int_0^T |S(x_{t-\tau(r, r')}) - S(y_{t-\tau(r, \tilde{r}')} )| dt d\xi((x, r'), (y, \tilde{r}')) \right\} \\ (23) \quad & \leq C_T \sqrt{1 + \chi(\mu)} d_T(\mu, \nu) \end{aligned}$$

by similar arguments.

We thus proved that, for all  $\mu, \nu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ :

$$(24) \quad |\Gamma_{1,\nu}(\mu) - \Gamma_1(\mu)| \leq C_T \sqrt{1 + \chi(\mu)} d_T(\mu, \nu).$$

### Proof of Lemma 12.(2.b)

Note that if  $\mu \not\ll P$  the inequality is trivial, since  $I(\mu|P) = \infty$ . Let then  $\mu \ll P$ . In order to obtain the proper inequality, we split it into four different terms. Let

us first remark that

$$\begin{aligned} |\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)| &\leq \frac{1}{2} \left| \int \int \left( \int G_t(r)(dW_t(x,r) - m_\nu(t,r)dt) \right)^2 d(\gamma_{\tilde{K}_{\nu,r}^T} - \gamma_{\tilde{K}_{\mu,r}^T}) d\mu(x,r) \right| \\ &+ \frac{1}{2} \left| \int \int \left\{ \left( \int G_t(r)(dW_t(x,r) - m_\mu(t,r)dt) \right)^2 - \left( \int G_t(r)(dW_t(x,r) - m_\nu(t,r)dt) \right)^2 \right\} d\gamma_{\tilde{K}_{\mu,r}^T} d\mu(x,r) \right| \\ &+ \left| \int \int (m_\nu - m_\mu)(t,r) dW_t(x,r) d\mu(x,r) \right|. \end{aligned}$$

Letting  $\xi$  be a probability measure on  $(\mathcal{C} \times D) \times (\mathcal{C} \times D)$  with marginals  $\mu$  and  $\nu$ , and  $\gamma_\xi$  be the law of a bidimensional centered Gaussian process  $(G, G')$  with covariance  $K_\xi$  (see (20)), we define

$$\Lambda_T(G(r)) = \frac{\exp\left(-\frac{1}{2} \int_0^T G_t(r)^2 dt\right)}{\int \exp\left(-\frac{1}{2} \int_0^T G_t(r)^2 dt\right) d\gamma_\xi}.$$

We then obtain:

$$\begin{aligned} (25) \quad |\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)| &\stackrel{C.S.}{\leq} \underbrace{\frac{1}{2} \int \int \left| \Lambda_T(G(r)) - \Lambda_T(G'(r)) \right| \left( \int G_t(r)(dW_t(x,r) - m_\nu(t,r)dt) \right)^2 d\gamma_\xi d\mu(x,r)}_{B_1} \\ &+ \underbrace{\frac{1}{2} \prod_{\varepsilon=\pm 1} \left( \int \int \Lambda_T(G'(r)) \left( \int (G_t(r) + \varepsilon G'_t)(dW_t(x,r) - m_\nu(t,r)dt) \right)^2 d\gamma_\xi d\mu \right)^{\frac{1}{2}}}_{B_2} \\ &+ \frac{1}{2} \underbrace{\left| \int \int \Lambda_T(G(r)) \left\{ \left( \int G_t(r)(dW_t(x,r) - m_\mu(t,r)dt) \right)^2 - \left( \int G_t(r)(dW_t(x,r) - m_\nu(t,r)dt) \right)^2 \right\} d\gamma_\xi d\mu \right|}_{B_3} \\ &\quad + \underbrace{\left( \int \left| \int (m_\nu - m_\mu)(t,r) dW_t(x,r) \right|^2 d\mu \right)^{\frac{1}{2}}}_{B_4}. \end{aligned}$$

Before bounding these four terms, we prove a useful inequality. For any  $h, m \in L^2([0; T], dt)$ , with  $m$  bounded:

$$(26) \quad \int \left( \int_0^T h_t(dW_t(x,r) - m(t)dt) \right)^2 d\mu_r(x) \leq 2 \left\{ \int \left( \int_0^T h_t dW_t(x,r) \right)^2 + \left( \int_0^T h_t m_t dt \right)^2 d\mu_r(x) \right\}.$$

Suppose that  $h \neq 0_{L^2([0; T], dt)}$ , then  $\Phi_h(x) = \frac{\left( \int_0^T h_t dW_t(x,r) \right)^2}{4 \left( \int_0^T h_t^2 dt \right)}$  is a well-defined, positive and measurable function of the sigma-algebra  $\mathcal{B}(\mathcal{C})$ , so that resorting to (13) one obtains

$$\int_{\mathcal{C}} \Phi_h(x) d\mu_r(x) \leq I(\mu|P_r) + \log \int_{\mathcal{C}} \exp \Phi_h(x) dP_r(x).$$

As  $W(., r)$  is a Brownian-motion under  $P_r$ ,  $\Phi_h \sim \mathcal{N}(0, \frac{1}{4})^2$ , so that Gaussian calculus gives, for any  $C > 2$ :

$$\int_C \left( \int_0^T h_t dW_t(x, r) \right)^2 d\mu_r(x) \leq C(I(\mu|P_r) + 1) \left( \int_0^T h_t^2 dt \right)$$

Remark that this inequality obviously holds also when  $h = 0_{L^2([0;T], dt)}$ . Applying this result in (26) one eventually finds:

$$\begin{aligned} \int \left( \int_0^T h_t (dW_t(x, r) - m(t)dt) \right)^2 d\mu_r(x) &\leq 2 \left\{ \left( C(1 + I(\mu_r|P_r)) + m_\infty^2 T \right) \left( \int_0^T h_t^2 dt \right) \right. \\ (27) \qquad \qquad \qquad &\leq C'(1 + I(\mu_r|P_r)) \left( \int_0^T h_t^2 dt \right). \end{aligned}$$

With this result in mind, we now control the first term. We start by noting that

$$\Lambda_T(G(r)) = \exp \left\{ -\Gamma_1(\mu, r) - \frac{1}{2} \int_0^T m_\mu^2(t, r) dt - \frac{1}{2} \int_0^T G_t(r)^2 dt \right\},$$

so that Jensen inequality ensures that

$$\Lambda_T(G(r)) \leq \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\}.$$

Consequently,

$$\begin{aligned} |\Lambda_T(G(r)) - \Lambda_T(G'(r))| &\leq \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\} \left( \frac{1}{2} \int_0^T |G_t(r)^2 - G'_t(r)^2| dt \right. \\ &\quad \left. + \left| \Gamma_1(\mu, r) - \Gamma_1(\nu, r) + \frac{1}{2} \int (m_\mu^2 - m_\nu^2)(t, r) dt \right| \right). \end{aligned}$$

$$\begin{aligned} B_1 &\stackrel{(27)}{\leq} C_T(I(\mu_r|P_r) + 1) \left( \int \left( \int_0^T |\Lambda_T(G(r)) - \Lambda_T(G'(r))| G_t(r)^2 dt \right) d\gamma_\xi \right) \\ &\stackrel{(22)}{\leq} C_T \sqrt{1 + \chi(\mu)} (I(\mu_r|P_r) + 1) d_T(\mu, \nu). \end{aligned}$$

As already observed in (17), when  $\mu \ll P$ , it can be decomposed as follow:

$$d\mu(x, r) = d\mu_r(x) c_\mu(r) d\pi(r),$$

where  $\mu_r \in \mathcal{M}_1^+(\mathcal{C})$  and  $c_\mu$  is a measurable positive function satisfying  $\mu_r \ll P_r$  and  $\int_D c_\mu(r) d\pi(r) = 1$  respectively. Integrating on  $D$  thus yields:

$$\begin{aligned} \int_D I(\mu_r|P_r) c_\mu(r) d\pi(r) &= \int \int \log \left( \frac{\rho_\mu(x, r)}{c_\mu(r)} \right) d\mu_r(x) c_\mu(r) d\pi(r) \\ &= \int \log(\rho_\mu(x, r)) d\mu(x, r) - \int \log(c_\mu(r)) c_\mu(r) d\pi(r) \stackrel{\text{Jensen}}{\leq} I(\mu|P), \end{aligned}$$

so that

$$\int_{C \times D} B_1(x, r) d\mu(x, r) = \int_D \int_C B_1(x, r) d\mu_r(x) c_\mu(r) d\pi(r) \leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu|P)) d_T(\mu, \nu).$$



Similarly, there exists a constant  $c_T$  such that

$$\begin{aligned} B_2 &\leq \frac{1}{2} \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\} \prod_{\varepsilon=\pm 1} \left( \int c_T (1 + I(\mu_r | P_r)) \left( \int \int (G_t(r) + \varepsilon G'_t(r))^2 dt d\gamma_\xi \right) c_\mu(r) d\pi(r) \right)^{\frac{1}{2}} \\ &\leq \frac{\|\sigma\|_\infty \sqrt{T}}{\lambda_*} \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\} \left( c_T (1 + I(\mu | P)) \right)^{\frac{1}{2}} \left( \int c_T (1 + I(\mu_r | P_r)) \right. \\ &\quad \left. \times \left( \int \int (G_t(r) - G'_t(r))^2 dt d\gamma_\xi \right) c_\mu(r) d\pi(r) \right)^{\frac{1}{2}}. \end{aligned}$$

As done previously, one can show that,

$$\int \int (G_t(r) - G'_t(r))^2 dt d\gamma_\xi \leq C_T (1 + \chi(\mu)) d_T(\mu, \nu)^2,$$

so that

$$B_2 \leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu | P)) d_T(\mu, \nu).$$

To bound  $B_3$ , we first use Cauchy-Schwarz inequality:

$$\begin{aligned} B_3 &\leq \frac{1}{2} \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\} \prod_{\varepsilon=\pm 1} \left\{ \int_D \int \int \left| \int_0^T G_t(r) ((1 + \varepsilon) dW_t(x, r) \right. \right. \\ (28) \quad &\quad \left. \left. - (m_\nu(t, r) + \varepsilon m_\mu(t, r)) dt \right|^2 d\gamma_\xi d\mu_r c_\mu(r) d\pi(r) \right\}^{\frac{1}{2}}. \end{aligned}$$

Then, by Cauchy-Schwarz inequality one observes that

$$\left| \int_0^T G_t(r) (m_\mu(t, r) - m_\nu(t, r)) dt \right|^2 \leq \left( \int_0^T G_t^2(r) dt \right) \left( \int_0^T (m_\mu(t, r) - m_\nu(t, r))^2 dt \right).$$

Remark that

$$\int_0^T (m_\mu - m_\nu)^2(t, r) dt = \int_0^T \left( \int_{C \times D} \frac{J(r, r')}{\lambda(r)} S(x_{t-\tau(r, r')}) d(\mu - \nu)(x, r) \right)^2 dt,$$

so that applying the same techniques, one finds that

$$\int \left| \int_0^T G_t(r) (m_\mu(t, r) - m_\nu(t, r)) dt \right|^2 d\gamma_\xi \leq C_T (1 + \chi(\mu)) d_T(\mu, \nu)^2.$$

Moreover, (27) gives:

$$\int \left\{ \int_0^T 2G_t(r) (dW_t(x, r) - \frac{m_\mu(t, r) + m_\nu(t, r)}{2} dt) \right\}^2 d\mu_r \leq c_T (1 + I(\mu_r | P_r)) \int_0^T G_t^2(r) dt.$$

Using the last two inequalities in (28) gives:

$$B_3 \leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu | P))^{\frac{1}{2}} d_T(\mu, \nu) \leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu | P)) d_T(\mu, \nu)$$

as  $I(\cdot | P) \geq 0$ .

As of the last term, we have

$$\begin{aligned} B_4 &\leq \left( \int_D c_T (1 + I(\mu_r | P_r)) \int_0^T (m_\mu(t, r) - m_\nu(t, r))^2 dt c_\mu(r) d\pi(r) \right)^{\frac{1}{2}} \\ &\leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu | P))^{\frac{1}{2}} d_T(\mu, \nu) \leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu | P)) d_T(\mu, \nu). \end{aligned}$$

Hence, we conclude that exists a constant  $C_T$  satisfying

$$|\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)| \leq C_T \sqrt{1 + \chi(\mu)} (1 + I(\mu|P)) d_T(\mu, \nu).$$

**Proof of Lemma 12.(3):** We proceed exactly as in Lemma 5.(vi) [13], remarking that  $\mathcal{C} \times D$  is also a Polish space, and that boundedness of  $I(\mu_{p_m}|P)$  implies that of  $\chi(\mu_{p_m})$ .  $\square$

**3.2. Upper-bound and Tightness.** We have proved that  $H = I(\cdot|P) - \Gamma$  is a good rate function, and we now want to show that it is indeed associated with a LDP. We demonstrate here a weak LDP relying on an upper-bound inequality for compact subsets, and tightness of the family  $(Q^N)_N$ . To prove the first point, we take advantage of the full LDP followed by  $\hat{\mu}_N$  under  $(Q_\nu)^{\otimes N}$ , and have to control an error term. This is done by carefully investigating Hölder continuity of the trajectories. Time-dependent delays can be handled under a short-time hypothesis, whereas constant delays present no specific difficulty. The second point will rely on the exponential tightness of  $P^{\otimes N}$ . These proofs are inspired from those of Guionnet in a non-spatial spin-glass model [25].

**Theorem 13.** *For  $T < \frac{\lambda_*^2}{2\|\sigma\|_\infty^2}$  and any compact subset  $K$  of  $\mathcal{M}_1^+(\mathcal{C} \times D)$ ,*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in K) \leq -\inf_K H.$$

*This result holds for any  $T > 0$  when delays are uniform in space.*

*Proof.* Let  $\delta < 0$ . We can find an integer  $M$  and a family  $(\nu_i)_{1 \leq i \leq M}$  of  $\mathcal{M}_1^+(\mathcal{C} \times D)$  such that

$$K \subset \bigcup_{i=1}^M B(\nu_i, \delta),$$

where  $B(\nu_i, \delta) := \{\mu | d_T(\mu, \nu_i) < \delta\}$ . A very classical result (see e.g. [19, lemma 1.2.15]), ensures that

$$\limsup \frac{1}{N} \log Q^N(\hat{\mu}_N \in K) \leq \max_{1 \leq i \leq M} \limsup \frac{1}{N} \log Q^N(\hat{\mu}_N \in K \cap B(\nu_i, \delta)).$$

Lemma 8 yields:

$$\begin{aligned} Q^N(\hat{\mu}_N \in K \cap B(\nu, \delta)) &= \int_{\hat{\mu}_N \in K \cap B(\nu, \delta)} \exp \left\{ N \Gamma(\hat{\mu}_N) \right\} dP^{\otimes N} \\ &= \int_{\hat{\mu}_N \in K \cap B(\nu, \delta)} \exp \left\{ N (\Gamma(\hat{\mu}_N) - \Gamma_\nu(\hat{\mu}_N)) \right\} \exp \left\{ N \Gamma_\nu(\hat{\mu}_N) \right\} dP^{\otimes N}. \end{aligned}$$

Observe that, for conjugate exponents  $(p, q)$ ,

$$\begin{aligned} Q^N(\hat{\mu}_N \in K \cap B(\nu, \delta)) &= \int_{\hat{\mu}_N \in K \cap B(\nu, \delta)} \exp \left\{ N (\Gamma(\hat{\mu}_N) - \Gamma_\nu(\hat{\mu}_N)) \right\} dQ_\nu^{\otimes N} \\ &\leq Q_\nu^{\otimes N}(\hat{\mu}_N \in K \cap B(\nu, \delta))^{\frac{1}{p}} \left( \int_{\hat{\mu}_N \in K \cap B(\nu, \delta)} \exp \left\{ q N (\Gamma(\hat{\mu}_N) - \Gamma_\nu(\hat{\mu}_N)) \right\} dQ_\nu^{\otimes N} \right)^{\frac{1}{q}}. \end{aligned}$$

The first term of the right hand side can be controlled by large deviations estimates. In the case where delays do not depend on the space variable, the second term

presents no specific difficulty. It becomes much more complex in the case of space-dependent delays. To cope with it, we restrict the integral to a specific subset of  $(\mathcal{C} \times D)^N$ . Let  $\frac{1}{4} < \beta < \frac{1}{2}$ , and  $E_{N,\delta}^j := \left\{ \sup_{t,s \in [-\tau, T], |t-s| \leq \delta} |x_t^j - x_s^j| \leq \delta^\beta \right\}$ . Let also  $c_{N,\delta} := \sum_{j=1}^N \mathbf{1}_{(E_{N,\delta}^j)^c}$ . It is the number of indices  $j$  for which problems appear. Let  $A_{N,\delta}^1 = \{c_{N,\delta} \leq \delta^{2\beta} N\}$ . As observed in remark 4, we can alternatively link measure  $\mu$  to the Gaussian process  $G$  or to the measure  $\gamma$ . In the following, it is convenient to rely on the former formulation in order to use Fubini theorem, in the demonstration of the following Lemma.

Let, for every  $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ ,

$$X_i^\mu = \int_0^T (G_t^\mu(r_i) + m_\mu(t, r_i)) dW_t(x^i, r_i) - \frac{1}{2} \int_0^T (G_t^\mu(r_i) + m_\mu(t, r_i))^2 dt.$$

Then, using the same kind of arguments as in Lemma 8:

$$\begin{aligned} Q^N(\hat{\mu}_N \in K \cap B(\nu, \delta)) &= \int \mathcal{E}_J \left( \int \mathbf{1}_{\{\hat{\mu}_N \in K \cap B(\nu, \delta)\} \cap A_{N,\delta}^1} dQ_{\mathbf{r}}^N(J)(\mathbf{x}) \right) d\pi^{\otimes N}(\mathbf{r}) + Q^N(A_{N,\delta}^{1,c}) \\ &= \int \mathcal{E} \left( \mathbf{1}_{\{\hat{\mu}_N \in K \cap B(\nu, \delta)\} \cap A_{N,\delta}^1} \prod_{i=1}^N \exp \{X_i^{\hat{\mu}_N}\} \right) \exp \{-N\Gamma_\nu(\hat{\mu}_N)\} dQ_\nu^{\otimes N}(\mathbf{x}, \mathbf{r}) + Q^N(A_{N,\delta}^{1,c}) \\ &= \int \mathcal{E} \left[ \mathbf{1}_{\{\hat{\mu}_N \in B(\nu, \delta)\} \cap A_{N,\delta}^1} \prod_{i=1}^N \left( \exp(X_i^{\hat{\mu}_N} - X_i^\nu) \right) \frac{\prod_{i=1}^N \exp X_i^\nu}{\mathcal{E} \left[ \prod_{i=1}^N \exp X_i^\nu \right]} \right] dQ_\nu^{\otimes N} + Q^N(A_{N,\delta}^{1,c}) \\ (29) \quad &\stackrel{\text{H\"older}}{\leq} Q^N(A_{N,\delta}^{1,c}) + Q_\nu^{\otimes N}(\hat{\mu}_N \in K \cap B(\nu, \delta))^{\frac{1}{p}} \\ &\times \underbrace{\left\{ \int \mathcal{E} \left[ \mathbf{1}_{\{\hat{\mu}_N \in B(\nu, \delta)\} \cap A_{N,\delta}^1} \prod_{i=1}^N \left( \exp q(X_i^{\hat{\mu}_N} - X_i^\nu) \exp X_i^\nu \right) \right] dP^{\otimes N} \right\}^{\frac{1}{q}}}_{B_N}. \end{aligned}$$

In Appendix C, we prove the following.

**Lemma 14.** (1) *For any real number  $q > 1$ , there exists a strictly positive real number  $\delta_q$  such that, for any  $\delta < \delta_q$ , there exists a function  $C_q(\cdot)$  in  $\mathbb{R}$  such that  $\lim_{\delta \rightarrow 0} C_q(\delta) = 0$  and:*

$$B_N \leq \exp\{C_q(\delta)N\}.$$

*Moreover, the result holds, without restricting the integral to  $A_{N,\delta}^1$  for spatially uniform delays.*

$$(2) \text{ For } T < \frac{\lambda_s^2}{2\|\sigma\|_\infty^2}, \lim_{\delta \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \left\{ Q^N(A_{N,\delta}^{1,c}) \right\} = -\infty.$$

*Remark 7. Remark that  $A_{N,\delta}^1$  takes into account the trajectories for negative times up to  $-\tau$ . In order to control this set, we make the natural assumption that initial conditions are well-behaved (e.g. as independent semi-martingales with constant diffusion coefficient).*

Let us now conclude the proof when delays depend upon space. We recall that  $\hat{\mu}_N$  satisfies a full LDP under  $Q_\nu^{\otimes N}$ :

$$\forall A \subset \mathcal{M}_1^+(\mathcal{C} \times D), -\inf_{A^o} H_\nu \leq \liminf_{N \rightarrow \infty} \frac{1}{N} \log Q_\nu^{\otimes N}(\hat{\mu}_N \in A) \leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log Q_\nu^{\otimes N}(\hat{\mu}_N \in A) \leq -\inf_A H_\nu.$$

Then, taking  $\delta < \delta_q$ , we find

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in K \cap B(\nu, \delta)) &\leq \max \left\{ \limsup_{N \rightarrow \infty} \frac{1}{N} \left( \frac{1}{p} \log Q_\nu^{\otimes N}(\hat{\mu}_N \in K \cap B(\nu, \delta)) + \frac{1}{q} C_q(\delta) N \right); \right. \\ &\quad \left. \limsup_{N \rightarrow \infty} \frac{1}{N} \log \left\{ Q^N((A_{N,\delta}^1)^c) \right\} \right\}. \end{aligned}$$

Furthermore, for  $\delta < \delta_q$

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \left( \frac{1}{p} \log Q_\nu^{\otimes N}(\hat{\mu}_N \in K \cap B(\nu, \delta)) + \frac{1}{q} C_q(\delta) N \right) &\leq -\frac{1}{p} \inf_{K \cap B(\nu, \delta)} H_\nu + \frac{1}{q} C_q(\delta) \\ &\leq -\frac{1}{p} \inf_{1 \leq i \leq M} \inf_{K \cap B(\nu_i, \delta)} H_{\nu_i} + \frac{1}{q} C_q(\delta) \\ &\leq -\frac{1}{p} \inf_{1 \leq i \leq M} \inf_{K \cap B(\nu_i, \delta)} (I(|P|) - \Gamma) - \frac{1}{p} \inf_{1 \leq i \leq M} \inf_{K \cap B(\nu_i, \delta)} (\Gamma - \Gamma_{\nu_i}) + \frac{1}{q} C_q(\delta). \end{aligned}$$

As in Lemma 12.(2), we can prove that there exists a finite constant  $C$  such that:

$$|\Gamma_{\nu_i}(\mu) - \Gamma(\mu)| \leq C \sqrt{1 + \chi(\mu)} (I(\mu|P) + 1) d_T(\nu_i, \mu),$$

so that:

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in K) &\leq \max \left\{ -\frac{1}{p} \inf_K H - \frac{1}{p} \inf_K \left( \sqrt{1 + \chi(\mu)} (1 + I(\mu|P)) C \delta \right) + \frac{1}{q} C_q(\delta); \right. \\ &\quad \left. \limsup_{N \rightarrow \infty} \frac{1}{N} \log \left\{ Q^N((A_{N,\delta}^1)^c) \right\} \right\}. \end{aligned}$$

Suppose  $I(\cdot|P)$  is infinite everywhere on  $K$ . Then, proposition 12.(3) ensures that  $H$  is also uniformly infinite on this compact set, so that

$$-\frac{1}{p} \inf_K H - \frac{1}{p} \inf_K \left( \sqrt{1 + \chi(\mu)} (1 + I(\mu|P)) C \delta \right) + \frac{1}{q} C_q(\delta) = -\inf_K H = -\infty.$$

Now, if exists a probability  $\mu \in K$  such that  $I(\mu|P) < \infty$ . Then, both  $\chi$  and  $H$  are also finite by the same argument. In every cases, letting  $\delta \searrow 0$  yields

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in K) \leq -\frac{1}{p} \inf_K H.$$

One concludes by sending  $p \searrow 1$ .

Uniform delays are treated similarly.  $\square$

**Theorem 15** (Tightness). *For any real number  $\varepsilon > 0$ , there exists a compact set  $K_\varepsilon$  of  $\mathcal{M}_1^+(\mathcal{C} \times D)$  such that, for any integer  $N$ ,*

$$Q^N(\hat{\mu}_N \notin K_\varepsilon) \leq \varepsilon.$$

*Proof.* The proof of this theorem consists in using the relative entropy inequality (13) and the exponential tightness of the sequence  $(P^{\otimes N})_N$ . The reader shall refer to [13, Theorem 2], to obtain inequality:

$$\begin{aligned} I(Q^N | P^{\otimes N}) &= \frac{N}{2} \int_{D^N} \int_{C^N} \int_0^T \left( \int_0^t \tilde{K}_{\hat{\mu}_N, r_1}^t(t, s) (dW_s(x^1, r_1) - m_{\hat{\mu}_N}(s, r_1) ds) \right. \\ &\quad \left. + m_{\hat{\mu}_N}(t, r_1) \right)^2 dt dQ_{\mathbf{r}}^N(\mathbf{x}) d\pi^{\otimes N}(\mathbf{r}) \\ &\leq N \left\{ \underbrace{\int_{D^N} \int_0^T \int_{C^N} \left( \int_0^t \tilde{K}_{\hat{\mu}_N, r_1}^t(t, s) (dW_s(x^1, r_1) - m_{\hat{\mu}_N}(s, r_1) ds) \right)^2 dQ_{\mathbf{r}}^N(\mathbf{x}) dt d\pi^{\otimes N}(\mathbf{r})}_{\varphi(t, \mathbf{r})} \right. \\ &\quad \left. + \frac{\|\bar{J}\|_{\infty}^2 T}{\lambda_*^2} \right\}. \end{aligned}$$

We then bound  $\varphi(t, \mathbf{r})$  uniformly in space to conclude:

$$\sup_{t \leq T} \varphi(t, \mathbf{r}) \leq 2 \frac{\|\sigma\|_{\infty}^4 T}{\lambda_*^4} \exp \left\{ 2 \frac{\|\sigma\|_{\infty}^4 T}{\lambda_*^4} \right\}.$$

□

#### 4. IDENTIFICATION OF THE MEAN-FIELD EQUATIONS

In the Gaussian case, we have seen that the series of empirical measures  $(\hat{\mu}_N)_N$  satisfies a large deviations principle of speed  $N$ , and with good rate function  $H$ . In order to identify the limit of the system, we study the minima of the functions  $H$  through a variational study of the function  $H$ , as in non-spatial networks in [5, 13]. Here, very little difficulties arise from spatiality, so that we will mainly state the results and refer to the two above articles for the proofs. We will show that any minimum  $Q$  of  $H$  satisfies:

$$(30) \quad Q \simeq P, \quad \frac{dQ}{dP}(x, r) = \int \exp \left\{ \int_0^T G_t(r) + m_Q(t, r) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_Q(t, r))^2 dt \right\} d\gamma_Q.$$

We will then prove that there exists a unique probability measure satisfying (30).

**4.1. Variational characterization of the minima of  $H$ .** The large deviations principle ensures that the possible limits of the empirical measures minimize the good rate functions. We hence need to identify these minima and show that they are uniquely defined by equation (30). In this purpose, we start by showing that any probability minimizing  $H$  is equivalent to  $P$ :

**Lemma 16.** *Let  $Q$  be a probability measure on  $\mathcal{C} \times D$  which minimizes  $H$ . Then  $Q \ll P$ . Moreover, defining  $B := \{(x, r); \frac{dQ}{dP}(x, r) = 0\}$ ,  $\delta := P(B)$ , and  $Q_s := \frac{Q + s \mathbf{1}_B P}{1 + s\delta}$ , one has:*

- $I(Q_s | P) = I(Q | P) + s\delta \log(s) + O(s)$ ,
- $\Gamma(Q_s) = \Gamma(Q) + O(s)$ .

*Proof.* The proof is done exactly as in [13]. One just has to remark that the location  $r$  plays no role in the expansion. □

This lemma ensures that any minimum  $H$  is equivalent to  $P$ . Indeed, if  $\delta > 0$ , then the result of lemma 16 implies that  $H(Q_s) - H(Q) \sim s\delta \log(s)$  which is strictly negative for small  $s$ , contradicting minimality  $H$  of at  $Q$ . We therefore necessarily have  $\delta = P(B) = 0$ , hence  $P \ll Q$ , i.e.  $Q \simeq P$ .

Let us now characterize the minima of  $H$ . We use a variational formulation to show that any minimum of  $H$  satisfies equation (30), and start by proving the following lemma:

**Lemma 17.** *Let  $\Phi$  be a positive and bounded measurable function on  $\mathcal{C} \times D$  such that  $\int \Phi dQ = 1$ , and denote  $\Psi := \Phi - 1$  and  $Q_s(\Phi) := \frac{1+s\Phi}{1+s}Q$ . We have:*

- $I(Q_s(\Phi)|P) = I(Q|P) + s \int \Psi \log \frac{dQ}{dP} dQ + O(s^2)$ ,
- $\Gamma(Q_s(\Phi)) = \Gamma(Q) + s \left\{ \int \left( \log \int \exp \left\{ \int_0^T (G_t(r) + m_Q(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_Q(t, r))^2 dt \right\} d\gamma_Q + Y_T \right) \Psi dQ + \int Y_T(y, r) dQ(y, r) + C_Q(\Phi) \right\} + O(s^{\frac{3}{2}})$  where  $Y_T$  is an adapted process with finite variation and  $C_Q$  is a bounded function.

*Proof.* The proof proceeds exactly as in [13].  $\square$

We are now able to prove that any minimum satisfies equation (30). A necessary condition for  $Q$  to minimize  $H$  is,

$$(31) \quad \lim_{s \rightarrow 0} \frac{1}{s} (H(Q_s(\Phi)) - H(Q)) \geq 0.$$

Moreover, Lemma 17 implies that, for every  $\Phi$  satisfying  $\int \Phi dQ = 1$ ,

$$(32) \quad \begin{aligned} H(Q_s(\Phi)) - H(Q) &= s \int \left\{ \log \frac{dQ}{dP} - \log \left( \int \exp \left\{ \int_0^T G_t(r) + m_Q(t, r) dW_t(x, r) \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{1}{2} \int_0^T (G_t(r) + m_Q(t, r))^2 dt \right\} d\gamma_Q \right) - Y_T \right\} \Psi dQ - s \left\{ \int Y_T dQ + C_Q(\Phi) \right\} + O(s^{\frac{3}{2}}). \end{aligned}$$

Let

$$\begin{aligned} Z_T(x, r) &= \log \frac{dQ}{dP}(x, r) - \log \left( \int \exp \left\{ \int_0^T G_t(r) + m_Q(t, r) dW_t(x, r) \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \int_0^T (G_t(r) + m_Q(t, r))^2 dt \right\} d\gamma_Q \right) - Y_T(x, r). \end{aligned}$$

We can rewrite (32):

$$H(Q_s(\Phi)) - H(Q) = s \int Z_T(x, r) \Psi dQ(x, r) - s \left\{ \int Y_T(x, r) dQ(x, r) + C_Q(\Phi) \right\} + O(s^{\frac{3}{2}}).$$

For any bounded measurable function  $\Psi$ , with  $\int \Psi dQ = 0$ , we claim that  $\int \Psi Z_T dQ = 0$ . Indeed, if not, we can find  $\Psi$  such that  $\int \Psi Z_T dQ \neq 0$  and  $\int \Psi dQ = 0$ . Now let  $\Psi_c = c\Psi$ . Choosing  $c = -d \operatorname{sign}(\int \Psi Z_T dQ)$  with  $d > 0$  large enough, it is clear that  $\Psi_c$  satisfies the condition of Lemma 17. Yet:

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{1}{s} (H(Q_s(\Phi_c)) - H(Q)) &= c \int \Psi Z_T dQ - \int Y_T dQ - C_Q(\Phi_c) \\ &\leq -d \left| \int \Psi Z_T dQ \right| - \int Y_T dQ + C_Q^\infty, \end{aligned}$$

which is strictly negative for  $d$  big enough. Hence, we find a contradiction with condition (31), so that  $\int \Psi_Z dQ = 0$ . As this result holds for every bounded measurable function with  $Q$ -measure zero, we can find a constant  $c_Q$  such that  $Z_T = c_Q$  almost surely under  $Q$  (and also under  $P$  by equivalence of these measures). But  $\left(\frac{dQ}{dP} \Big|_{\mathcal{F}_t}\right)_{t \leq T}$  must be a  $(\mathcal{C} \times D, (\mathcal{F}_t)_{t \leq T}, \mathcal{F}_T, P)$ -local martingale. Since

$\left(\int \exp \left\{ \int_0^t G_s(r) + m_Q(s, r) dW_s(x, r) - \frac{1}{2} \int_0^t (G_s(r) + m_Q(s, r))^2 ds \right\} d\gamma_Q \right)_{t \leq T}$  is a  $P$ -local martingale and  $(Y_t)_{t \leq T}$  is a process with finite variation, one has by uniqueness of semimartingale decomposition,

$$\frac{dQ}{dP} = \int \exp \left\{ \int_0^T G_t(r) + m_Q(t, r) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_Q(t, r))^2 dt \right\} d\gamma_Q.$$

As a consequence of the form of the density obtained from equation (30), the characterization of the minima readily implies (as a simple application of Girsanov theorem) that the possible limits of the network equations are the law of the solutions of the implicit equation:

$$(33) \quad d\bar{X}_t(r) = \left( f(r, t, \bar{X}_t(r)) + U_t^{\bar{X}}(r) \right) dt + \lambda(r) dW_t(r)$$

where the processes  $r \rightarrow (W_t(r))$  are independent Brownian motions and the processes  $r \rightarrow (U_t^{\bar{X}}(r))$  are independent Gaussian processes with mean

$$m(t, r) = \int_D J(r, r') \mathbb{E}_Z \left[ S(\bar{Z}_{t-\tau(r, r')}(r')) \right] d\pi(r')$$

and covariance

$$C(t, s, r) = \int_D \sigma(r, r')^2 \mathbb{E}_Z \left[ S(\bar{Z}_{s-\tau(r, r')}(r')) S(\bar{Z}_{t-\tau(r, r')}(r')) \right] d\pi(r'),$$

where  $\bar{Z}$  denotes an independent copy of  $\bar{X}$ , and  $\mathbb{E}_Z$  is the expectation over  $\bar{Z}$ . Hence, we exactly find the mean-field limit given by the heuristic approach summarized in section 1.2.

**4.2. Uniqueness of the minimum.** In this section, we rely on a contraction argument to show that the good rate function  $H$  admits a unique minimum. The proof roughly follows the steps of the spin-glass non-spatialized problem[5, Theorem 5.5], but several technicalities appear due to the spatial nature of the system and specificities of the neuronal setting. We provide here details of the proof which was yet to be developed in the context of the neuronal equations<sup>1</sup>.

We start by introducing the map:

$$F_\mu := \begin{cases} \mathcal{C} \times D \rightarrow \mathbb{R} \\ (x, r) \rightarrow \int \exp \left\{ \int_0^T (G_t(r) + m_\mu(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_\mu(t, r))^2 dt \right\} d\gamma_\mu. \end{cases}$$

This is a non-negative measurable function for every  $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ , so that the map

$$L := \begin{cases} \mathcal{M}_1^+(\mathcal{C} \times D) \rightarrow \mathcal{M}_1^+(\mathcal{C} \times D) \\ \mu \rightarrow L(\mu) \end{cases}$$

---

<sup>1</sup>For instance, in the non-spatialized case treated in [13] was used a strong assumption of linearity of the intrinsic dynamics (our function  $f$ ), which implied that solutions were Gaussian and moment methods were used (see [22]).



defined by  $\frac{dL(\mu)}{dP}(x, r) := F_\mu(x, r)$  is a positive measure on  $\mathcal{C} \times D$ . In fact, remark that  $F_\mu$  is a positive continuous function of  $(x, r)$ , as a continuous function of  $m_\mu(\cdot, r)$ ,  $K_\mu(\cdot, \cdot, r)$  and  $W(x, r)$  which are both continuous maps. Moreover, as  $\exp \left\{ \int_0^T (G_t(r) + m_\mu(t, r)) dW_t(x, r) - \frac{1}{2} \int_0^T (G_t(r) + m_\mu(t, r))^2 dt \right\}$  is  $\gamma_\mu$ -almost surely finite, one can use Novikov criterion to show that  $L(\mu)$  defines a probability measure on  $\mathcal{C} \times D$ . The main result of the section is the following:

For every  $C > 0$ , we introduce the set

$$\mathcal{A}_{T,C} := \left\{ \mu \in \mathcal{M}_1^+(\mathcal{C} \times D) \mid \mu \ll P, \forall r \in D \int_{\mathcal{C}} \frac{d\mu}{dP}(x, r) dP_r(x) = 1, \chi(\mu) \leq C \right\}.$$

The first condition restrains the solution we are considering to those that are almost continuous with respect to  $P$ . This is a natural constraint in regards to the results obtained in the characterization of the solutions. The second condition states that if  $\mu \in \mathcal{A}_{T,C}$ , then it defines a proper probability measure at every location  $r \in D$ . In fact, if we let  $d\mu_r := \frac{d\mu}{dP}(\cdot, r) dP_r$ , then, it is clear that  $\mu_r \in \mathcal{M}_1^+(\mathcal{C})$ , and we have  $d\mu(x, r) = d\mu_r(x) d\pi(r)$ . The third condition states that at each point, the trajectories of our neurons must have some regularity in time (Hölder 1/2 in the sense defined previously). To summarize, we are only considering solutions that define proper stochastic processes with regular trajectories at each location on  $D$ .

**Theorem 18.** *There exists a universal constant  $C > 0$ , such that*

- (1)  $P \in \mathcal{A}_{T,C}$ ,
- (2)  $L(\mathcal{M}_1^+(\mathcal{C} \times D)) \subset \mathcal{A}_{T,C}$ . In particular  $\mathcal{A}_{T,C}$  is stable by  $L$ .
- (3) The map  $L$  admits a unique fixed point on  $\mathcal{A}_{T,C}$ .

*Proof.* (1)-(2):

Let  $\mathcal{F}_t = \sigma(x_s(r), s \leq t, r \in D)$  be the natural filtration on  $\mathcal{M}_1^+(\mathcal{C} \times D)$  and  $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$ . We want to show that  $L(\mu) \in \mathcal{A}_{T,C}$  for some  $C$  big enough. By definition,  $L(\mu) \ll P$ . As in [5, lemma 5.15] we can express the density of  $L(\mu)$  with respect to  $P$  as the exponential of a local martingale:

$$\forall t \leq T, \quad \frac{dL(\mu)}{dP} \Big|_{\mathcal{F}_t}(x, r) = \exp \left\{ \underbrace{\int_0^t H_s^\mu(x, r) dW_s(x, r) - \frac{1}{2} \int_0^t H_s^\mu(x, r)^2 ds}_{X_t^\mu(x, r)} \right\},$$

where  $H_t^\mu(x, r) = \int_0^t \tilde{K}_{\mu, r}^t(t, s) (dW_s(x, r) - m_\mu(s, r) ds) + m_\mu(t, r)$ . Remark that the non-centered Gaussian weights bring new terms in the expression of  $H_t^\mu(x, r)$ . As  $\frac{dL(\mu)}{dP}(\cdot, r) = \exp\{X_t^\mu(\cdot, r)\}$  is a  $P_r$ -martingale, then

$$\forall r \in D, \int_{\mathcal{C}} \frac{dL(\mu)}{dP}(x, r) dP_r(x) = 1,$$

so that the second condition is verified. Let us prove the following result that will expedite our analysis:

**Lemma 19.** *For  $p - 1 > 0$  small enough,  $\exists C_1 > 0, \forall r \in D, \forall \mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$*

$$\int \exp(p X_t^\mu(x, r)) dP_r(x) \leq C_1.$$

*Proof.* In fact

$$\begin{aligned}
\int \exp(pX_t^\mu(x, r)) dP_r(x) &= \int \left( \int \exp\left(\int_0^t G_s(r) + m_\mu(s, r) dW_s(x, r) \right. \right. \\
&\quad \left. \left. - \frac{1}{2} \int_0^t (G_s(r) + m_\mu(s, r))^2 ds\right) d\gamma_\mu \right)^p dP_r, \\
&\leq \int \int \exp\left(p \int_0^t G_s(r) + m_\mu(s, r) dW_s(x, r) - \frac{p}{2} \int_0^t (G_s(r) + m_\mu(s, r))^2 ds\right) d\gamma_\mu dP_r, \\
&= \int \exp\left(\frac{p(p-1)}{2} \int_0^t (G_s(r) + m_\mu(s, r))^2 ds\right) d\gamma_\mu,
\end{aligned}$$

where we used the martingale property. The right-hand side of the last inequality can be upperbounded by a uniform constant in space using (15).

One can also remark that, by Jensen inequality and Fubini theorem

$$\int \exp\left(\frac{p(p-1)}{2} \int_0^t (G_s(r) + m_\mu(s, r))^2 ds\right) d\gamma_\mu \leq \int_0^t \int \exp\left(\frac{p(p-1)t}{2} (G_s(r) + m_\mu(s, r))^2\right) d\gamma_\mu \frac{ds}{t}.$$

Using equation (15), and remarking that  $K_\mu(s, s, r) \leq \frac{\|\sigma\|_\infty^2}{\lambda_*^2}$  and  $m_\mu(s, r)^2 \leq \frac{\|\bar{J}\|_\infty^2}{\lambda_*^2}$ , one obtains choosing  $p-1$  small enough,

$$\int \exp\left(\frac{p(p-1)}{2} \int_0^t (G_s(r) + m_\mu(s, r))^2 ds\right) d\gamma_\mu \leq \exp\left\{(p-1) \frac{T\|\sigma\|_\infty^2}{2\lambda_*^2} + \underbrace{o(p-1)}_{\text{uniform in } r}\right\}.$$

□

Now,  $\forall s < t \in [-\tau, T]$ , and any conjugate exponent  $(p, q)$  with  $p-1 > 0$  small enough, one has

$$\int_{\mathcal{C} \times D} |x_t - x_s|^2 dL(\mu)(x, r) \leq \left( \int_{\mathcal{C} \times D} |x_t - x_s|^{2q} dP(x, r) \right)^{\frac{1}{q}} C_1^{\frac{1}{p}}.$$

Now, fixing  $r \in D$ , one has

$$\begin{aligned}
\int_{\mathcal{C}} |x_t - x_s|^{2q} dP_r(x) &\leq c_q \left\{ \int_{\mathcal{C}} \left( \int_s^t |f(r, u, x_u)| du \right)^{2q} dP_r(x) + \int_{\mathcal{C}} \left( \lambda(r) (W_t(x, r) - W_s(x, r)) \right)^{2q} dP_r(x) \right\} \\
&\leq c_q \left\{ \int_{\mathcal{C}} (t-s)^{2q} \sup_{t \in [-\tau, T]} |f(r, t, x_t)|^{2q} dP_r(x) + \int_{\mathcal{C}} (\lambda^*)^{2q} (t-s)^q \left( \frac{W_t(x, r) - W_s(x, r)}{\sqrt{t-s}} \right)^{2q} dP_r(x) \right\} \\
&\leq C_q (t-s)^q \left( (t-s)^q E_{P_r} \left[ \sup_{t \in [-\tau, T]} |f(r, t, x_t)|^{2q} \right] + E[|\mathcal{N}(0, 1)|^{2q}] \right).
\end{aligned}$$

Since  $\sup_{t \in [-\tau, T]} |f(r, t, x_t)|$  admits moments of every polynomial order (see Appendix A, equation (43)). we have for  $q$  big enough

$$\chi(L(\mu)) \leq C_q^{\frac{1}{q}} \left( (T+\tau)^q \int_{\mathcal{C} \times D} \sup_{t \in [-\tau, T]} |f(r, t, x_t)|^{2q} dP(x, r) + E[|\mathcal{N}(0, 1)|^{2q}] \right)^{\frac{1}{q}} C_1^{\frac{1}{p}} = \tilde{C}_q.$$

The right-hand side is a universal constant independent of  $\mu$ . Moreover, it is easy to see that there exists a constant  $C_0 > 0$  such that  $P \in \mathcal{A}_{T, C_0}$ . We then take  $C = \max(C_0, \tilde{C}_q)$ .

(3):

Let  $P_t$  be the restriction of  $P$  to the  $\sigma$ -algebra  $\mathcal{F}_t$  (so that  $P = P_T$ ), and let  $D_T$  be the variational distance on  $\mathcal{M}_1^+(\mathcal{C} \times D)$  defined by

$$\forall \mu, \nu \in \mathcal{M}_1^+(\mathcal{C} \times D), \quad D_T(\mu, \nu) := \sup \left| \int f d\mu - \int f d\nu \right|,$$

where the supremum is taken on the measurable functions of  $\mathcal{C} \times D$  uniformly bounded by 1. On  $\mathcal{A}_{T,C}$ , one has

$$\forall \mu, \nu \in \mathcal{A}_{T,C}, \quad D_T(\mu, \nu) = \int \left| \frac{d\mu}{dP} - \frac{d\nu}{dP} \right| dP = \int \left| \frac{d\mu_r}{dP_r} - \frac{d\nu_r}{dP_r} \right| dP.$$

We will use the notations  $D_t(\mu, \nu) := D_T(\mu|_{\mathcal{F}_t}, \nu|_{\mathcal{F}_t})$ ,  $D_t^r(\mu, \nu) := \int \left| \frac{d\mu_r}{dP_r} - \frac{d\nu_r}{dP_r} \right|_{\mathcal{F}_t} dP_r$ . For every  $0 \leq t \leq T$ , we introduce the Vasertein-like distance on  $\mathcal{M}_1^+(\mathcal{C} \times D)$  which will naturally appear in our demonstration:

$$\forall \mu, \nu \in \mathcal{M}_1^+(\mathcal{C} \times D), \quad d_t^*(\mu, \nu) = \inf_{\xi} \left\{ \int \left( \sup_{-\tau \leq s \leq t} |x_s - y_s| + \sqrt{|r - r'|} \right) \wedge 1 \, d\xi((x, r), (y, r')) \right\}$$

where the infimum is taken on element of  $\mathcal{M}_1^+((\mathcal{C} \times D)^2)$  with marginal  $\mu$  and  $\nu$ .

It has the good property of being dominated by the variational distance. In detail, for all  $\mu, \nu \in \mathcal{A}_{T,C}$ ,  $r \in D$ , let  $\xi_r$  be a probability measure of  $\mathcal{M}_1^+(\mathcal{C}^2)$  with marginals  $\mu_r$  and  $\nu_r$ , and  $\tilde{\xi} \in \mathcal{M}_1^+((\mathcal{C} \times D)^2)$  such that  $d\tilde{\xi}((x, r), (y, r')) = d\xi_r(x, y) d\delta_r(r') d\pi(r)$ . In particular, it has marginals  $\mu$  and  $\nu$ , so that

$$\begin{aligned} d_t^*(\mu, \nu) &\leq \int_{(\mathcal{C} \times D)^2} \left( \sup_{-\tau \leq s \leq t} |x_s - y_s| + \sqrt{|r - r'|} \right) \wedge 1 \, d\tilde{\xi}((x, r), (y, r')) \\ &\leq \int_D \int_{\mathcal{C}^2} \sup_{-\tau \leq s \leq t} |x_s - y_s| \wedge 1 \, d\xi_r(x, y) d\pi(r). \end{aligned}$$

One can always chose distributions  $(\xi_r)_{r \in D}$  for the quantities  $\left( \int_{\mathcal{C}^2} \sup_{-\tau \leq s \leq t} |x_s - y_s| \wedge 1 d\xi_r(x, y) \right)_{r \in D}$  to be arbitrarily close to their infima. Moreover, it is well known that Vasertein-1 distance (Kantorovich-Rubinstein) on  $\mathcal{M}_1^+(\mathcal{C})$  is dominated by the variational distance. Hence, for  $\mu, \nu \in \mathcal{A}_{T,C}$

$$(34) \quad d_t^*(\mu, \nu) \leq \int_D D_t^r(\mu, \nu) d\pi(r).$$

It presents no difficulty to use the proof of [5, Lemma 5.15] to show that for any conjugate exponents  $(p, q)$  and any measure in  $\mathcal{A}_{T,C}$ :

$$\begin{aligned} D_t^r(L(\mu), L(\nu)) &\leq \left( \int |X_t^\mu(x, r) - X_t^\nu(x, r)|^q dP_r \right)^{\frac{1}{q}} \\ &\times \left( \int_0^1 \left( \int \exp \{p X_t^\mu(x, r)\} dP_r \right)^{1-\alpha} \left( \int \exp \{p X_t^\nu(x, r)\} dP_r \right)^\alpha d\alpha \right)^{\frac{1}{p}}. \end{aligned}$$

The second term of the product is controlled using lemma 19, whereas the first one is by:

**Lemma 20.**  $\forall q \geq 2, \exists c_q, C_T > 0$  depending respectively only on  $q$  and  $T$ , such that

$$(35) \quad \forall \mu, \nu \in \mathcal{A}_T, \forall r \in D, \int |X_t^\mu(x, r) - X_t^\nu(x, r)|^q dP_r \leq c_q C_T^q \left( \int_0^t D_s^r(\mu, \nu)^q ds + \int_0^t d_s^*(\mu, \nu)^q ds \right).$$

*Proof.* Burkholder-Davis-Gundy (BDG) inequality yields:

$$\begin{aligned} & \int \left| \int_0^t (H_s^\mu(x, r) - H_s^\nu(x, r)) dW_s(x, r) \right|^q dP_r(x) \\ & \stackrel{\text{BDG}}{\leq} c_q \int \left( \int_0^t (H_s^\mu(x, r) - H_s^\nu(x, r))^2 ds \right)^{\frac{q}{2}} dP_r(x) \\ & \stackrel{\text{Jensen, Fubini}}{\leq} c_q T^{\frac{q}{2}-1} \int_0^t \int |H_s^\mu(x, r) - H_s^\nu(x, r)|^q dP_r(x) ds. \end{aligned}$$

Moreover,

$$\begin{aligned} H_s^\mu(x, r) - H_s^\nu(x, r) &= \underbrace{\int_0^s (\tilde{K}_{\mu,r}^s(s, u) - \tilde{K}_{\nu,r}^s(s, u)) dW_u(x, r)}_{A_1} + \underbrace{\int_0^s (\tilde{K}_{\mu,r}^s(s, u) - \tilde{K}_{\nu,r}^s(s, u)) m_\mu(u, r) du}_{A_2} \\ &+ \underbrace{\int_0^s \tilde{K}_{\nu,r}^s(s, u) (m_\mu(u, r) - m_\nu(u, r)) du}_{A_3} + \underbrace{m_\mu(s, r) - m_\nu(s, r)}_{A_4} \end{aligned}$$

As  $q \geq 2$ , exists a constant

$$\int |H_s^\mu(x, r) - H_s^\nu(x, r)|^q dP_r \leq c_q \sum_{i=1}^4 \int |A_i|^q dP_r.$$

To treat the terms  $A_3$  and  $A_4$ , we refine (23):

$$\begin{aligned} |(m_\mu - m_\nu)|(t, r) &\leq \frac{1}{\lambda_*} \int |J(r, r') S(x_{t-\tau(r, r')}) - J(r, \tilde{r}') S(y_{t-\tau(r, \tilde{r}')})| d\xi((x, r'), (y, \tilde{r}')) \\ &\leq \frac{1}{\lambda_*} \left\{ \int (K_J |\tilde{r}' - r'|) \wedge (2 \|\bar{J}\|_\infty) d\xi + \|\bar{J}\|_\infty \int |S(x_{t-\tau(r, r')}) - S(y_{t-\tau(r, \tilde{r}')})| d\xi \right\} \\ &\leq C \left\{ \int |\tilde{r}' - r'| \wedge 1 d\xi + \int \left( \sup_{s \in [-\tau, t]} |x_s - y_s| \right) \wedge 1 d\xi + \int |S(x_{t-\tau(r, r')}) - S(x_{t-\tau(r, \tilde{r}')})| d\xi \right\} \\ &\leq C \left\{ \int d_D^{\frac{1}{2}} |\tilde{r}' - r'|^{\frac{1}{2}} \wedge 1 d\xi + \int \left( \sup_{s \in [-\tau, t]} |x_s - y_s| \right) \wedge 1 d\xi + \chi(\mu) \int \sqrt{\frac{|\tilde{r}' - r'|}{c}} \wedge \sqrt{2\tau} d\xi \right\}. \end{aligned}$$

Hence,  $|(m_\mu - m_\nu)|(t, r) \leq C d_t^*(\mu, \nu)$  with  $C$  uniform in space. Moreover, uniform boundedness of  $|\tilde{K}_{\mu,r}^s(s, u)| \leq \frac{\sigma^2}{\lambda^2} \exp\left(\frac{T\sigma^2}{2\lambda^2}\right)$  brings that  $A_3, A_4 \leq C_T d_s^*(\mu, \nu)$  with  $C_T$  uniform in  $D$ . To handle the term  $A_2$  we recall that for fixed  $r \in D$ , it is proven in [5] Lemma A.4 that the function  $\mu \rightarrow \tilde{K}_{\mu,r}^s$  is uniformly Lipschitz with respect to the Kantorovich-Rubinstein distance. In [5], the latter metric is slightly distinct and does not necessitate bounding the integrand as we do here, taking advantage of the boundedness of their processes. Here, this regularity clearly persists, because of the uniform boundedness and Lipschitz continuity of the functions  $S, J, \sigma$ . Moreover the uniform bounds in space ensure that the Lipschitz constant of  $\mu \rightarrow \tilde{K}_{\mu,r}^s$  can be chosen independent of  $r$ . Hence  $A_2 \leq C_T D_s^r(\mu, \nu)$ .

Let us eventually handle the term  $A_1$ . Using again Burkholder-Davis-Gundy and Jensen inequalities, one obtains

$$\int |A_1|^q dP_r \leq c_q T^{\frac{q}{2}-1} \int \int_0^s (\tilde{K}_{\mu,r}^s(s,u) - \tilde{K}_{\nu,r}^s(s,u))^q du dP_r \leq c_q C_T^q D_s^r(\mu, \nu)^q.$$

Let us now investigate the quadratic term. Let  $\tilde{H}_t^\mu(x, r) := H_t^\mu(x, r) - m_\mu(t, r)$ . Then

$$(36) \quad (H^\mu)^2 - (H^\nu)^2 = \underbrace{(\tilde{H}^\mu)^2 - (\tilde{H}^\nu)^2}_{\tilde{A}_1} + 2 \underbrace{(\tilde{H}^\mu - \tilde{H}^\nu)m_\mu}_{\tilde{A}_2} + 2 \underbrace{\tilde{H}^\nu(m_\mu - m_\nu)}_{\tilde{A}_3} + \underbrace{m_\mu^2 - m_\nu^2}_{\tilde{A}_4}.$$

As for the linear term, exists a constant  $c_q$  such that

$$\int \left| \int (H^\mu)^2 - (H^\nu)^2 dW \right|^q dP_r \leq c_q \sum_{i=1}^4 \int_0^t \int |\tilde{A}_i|^q dP_r ds.$$

The terms  $\tilde{A}_2$  and  $\tilde{A}_4$  can be bounded exactly as previously. The first one is the most delicate. We will control it last. Let us start by taking care of  $\tilde{A}_3$ .

Remark that

$$\left| \tilde{H}_s^\nu(x, r)(m_\mu(s, r) - m_\nu(s, r)) \right|^q \leq c_q C_T^q \left( \left| \int_0^s \tilde{K}_{\mu,r}^s(s, u) dW_u(x, r) \right|^q + T^q \|\tilde{K}\|_\infty^q \|m\|_\infty^q \right) d_s^*(\mu, \nu)^q.$$

Hence

$$\begin{aligned} \int |\tilde{A}_3|^q dP_r &\stackrel{BDG}{\leq} \tilde{c}_q C_T^q T^{\frac{q}{2}-1} \left( \int \left| \int_0^s \tilde{K}_{\mu,r}^s(s, u)^2 du \right|^{\frac{q}{2}} dP_r + 1 \right) d_s^*(\mu, \nu)^q \\ &\leq \tilde{c}_q C_T^q d_s^*(\mu, \nu)^q. \end{aligned}$$

To control the first term of (36), we use Itô formula to give a more suitable expression of  $(\tilde{H}^\mu)^2$ :

$$(\tilde{H}_s^\mu)^2 = 2 \int_0^s \tilde{H}_u^\mu \tilde{K}_{\mu,r}^s(u, u) dW_u + 2 \int \tilde{K}_{\mu,r}^s(u, u) \tilde{H}_u^\mu m_\mu(u, r) du + \int_0^s \tilde{K}_{\mu,r}^s(u, u)^2 du.$$

Then

$$\begin{aligned} (\tilde{H}_s^\mu)^2 - (\tilde{H}_s^\nu)^2 &= 2 \int_0^s (\tilde{H}_u^\mu - \tilde{H}_u^\nu) \tilde{K}_{\mu,r}^s(u, u) dW_u + 2 \int_0^s \tilde{H}_u^\nu (\tilde{K}_{\mu,r}^s(u, u) - \tilde{K}_{\nu,r}^s(u, u)) dW_u \\ &\quad + \int_0^s (\tilde{K}_{\mu,r}^s(u, u) - \tilde{K}_{\nu,r}^s(u, u)) (\tilde{K}_{\mu,r}^s(u, u) + \tilde{K}_{\nu,r}^s(u, u)) du + 2 \int_0^s (\tilde{H}_u^\nu - \tilde{H}_u^\mu) \tilde{K}_{\nu,r}^s(u, u) m_\nu(u, r) du \\ &\quad + \int_0^s \tilde{H}_u^\mu (\tilde{K}_{\nu,r}^s(u, u) - \tilde{K}_{\mu,r}^s(u, u)) m_\nu(u, r) du + \int_0^s \tilde{H}_u^\mu \tilde{K}_{\mu,r}^s(u, u) (m_\nu(u, r) - m_\mu(u, r)) du. \end{aligned}$$

After a fastidious calculus that presents no specific difficulty. We can bound each term of the previous equality. Gathering all estimates, we find:

$$\int |X_t^\mu(x, r) - X_t^\nu(x, r)|^q dP_r \leq c_q C_T^q \left( \int_0^t D_s^r(\mu, \nu)^q ds + \int_0^t d_s^*(\mu, \nu)^q ds \right).$$

□

We combine the result of the two lemmas as well as inequality (34), to obtain by Hölder inequality:

$$\forall \mu, \nu \in \mathcal{A}_{T,C}, \quad D_t^r(L(\mu), L(\nu))^q \leq c_q \tilde{C}_T^q \left( \int_0^t D_s^r(\mu, \nu)^q ds + \int_0^t \int_D D_s^r(\mu, \nu)^q d\pi(r) ds \right).$$

We can now integrate on space to find

$$\int_D D_t^r(L(\mu), L(\nu))^q d\pi(r) \leq c_q \tilde{C}_T^q \int_0^t \int_D D_s^r(\mu, \nu)^q d\pi(r) ds.$$

It is easy to see that  $\forall q \geq 1, \forall \mu, \nu \in \mathcal{A}_{T,C}$ , the application  $D_{t,q}(\mu, \nu) := \left( \int D_t^r(\mu, \nu)^q d\pi(r) \right)^{\frac{1}{q}}$  is a distance on  $\mathcal{A}_{T,C}$ . Choosing proper conjugate exponents  $(p, q)$ , we thus obtain, for every  $\mu, \nu \in \mathcal{A}_{T,C}$ :

$$D_{t,q}(L(\mu), L(\nu))^q \leq \lambda \int_0^t D_{s,q}(\mu, \nu)^q ds.$$

In particular, if one setting  $\mu_n := L^n(P)$ ,

$$D_{T,q}(\mu_{n+1}, \mu_n)^q \leq \lambda^n \int_{0 \leq s_1 \leq \dots \leq s_n \leq T} D_{s_1,q}(L(P), P)^q \prod_{i=1}^n ds_i \leq 2^q \frac{(\lambda T)^n}{n!},$$

eventually yielding

$$d_T(\mu_{n+1}, \mu_n) \leq 2 \left( \frac{(\lambda T)^n}{n!} \right)^{\frac{1}{q}}.$$

The conclusion is now classical. □

**4.3. Convergence of the process.** We are now in a position to prove theorem 3.

*Proof.* Indeed, for  $\delta$  a strictly positive real number and  $B(Q, \delta)$  the open ball of radius  $\delta$  centered in  $Q$  for the Vaserstein distance. We prove that  $Q^N(\hat{\mu}_N \notin B(Q, \delta))$  tends to zero as  $N$  goes to infinity. Indeed, for  $K_\varepsilon$  a compact defined in theorem 15, we have for any  $\varepsilon > 0$ :

$$Q^N(\hat{\mu}_N \notin B(Q, \delta)) \leq \varepsilon + Q^N(\hat{\mu}_N \in B(Q, \delta)^c \cap K_\varepsilon).$$

The set  $B(Q, \delta)^c \cap K_\varepsilon$  is a compact, and theorem 13 now ensures that

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in B(Q, \delta)^c \cap K_\varepsilon) \leq - \inf_{B(Q, \delta)^c \cap K_\varepsilon} H$$

and eventually, theorem 6 ensures that the righthand side of the inequality is strictly negative, which implies that

$$\lim_{N \rightarrow \infty} Q^N(\hat{\mu}_N \notin B(Q, \delta)) \leq \varepsilon,$$

that is:

$$\lim_{N \rightarrow \infty} Q^N(\hat{\mu}_N \notin B(Q, \delta)) = 0. \quad \square$$

## 5. NON GAUSSIAN CONNECTIVITY WEIGHTS

In this section, we relax the hypothesis that the synaptic weights are Gaussian. We only consider that the  $J_{ij}$  are i.i.d. random variables with sub-Gaussian tails (condition  $(H_J)$ ), mean  $\frac{J(r_i, r_j)}{N}$  and variance  $\frac{\sigma(r_i, r_j)^2}{N}$ . For technical reasons, we also assume here that the map  $\sigma$  is bounded away from zero:  $\exists \sigma_* > 0, \sigma(r, r') \geq \sigma_*$ . In this new setting, the LDP upper-bound of Theorem 4 no longer holds, its proof made important use of Gaussian properties. We do not extend here the LDP, but show that the empirical measure still converges towards the same process as in the Gaussian case, the unique minimum of the good rate function  $H$ .

We revisit technical tools developed by Moynot and Samuelides in [32] where they demonstrate similar results in a discrete time, non-spatialized setting. Their central idea is to show that the non-Gaussian and Gaussian density are exponentially close to one another, so that their quotient can be controlled by the exponential convergence of the Gaussian empirical measure toward  $Q$ . In that purpose, working in finite discretization of the time interval is mandatory, as the approach cannot be readily applied to a continuous-time settings. Technically, the estimates in [32] contain a sum over all the partition's times of squares of standard centered Gaussian variables, which would diverge with the discretization step going to zero. An additional error term - comparing continuous and discrete Gaussian densities - arises from the discretization which we need to control. Nonetheless, we will show that, under a short-time hypothesis and when the partition is fine enough, the error becomes uniformly controllable. In all the demonstration, it is of crucial importance to track down the effect of the size of the partition in every constant obtained for our upperbounds.

Let us introduce our regular discretization of time: we choose  $\frac{1}{\delta} \in \mathbb{N}^*$ , and define  $\Delta_\delta := \{t_l = l\delta T, l \in \llbracket 0, \frac{1}{\delta} \rrbracket\}$ , a regular partition of  $[0, T]$ . We denote by  $Q^{N, \delta}$  the solution of the SDE with non-Gaussian independent connectivity coefficients  $J_{ij}$

$$\begin{cases} dX_t^{i, N} = \left( f(r_i, t, X_t^{i, N}) + \sum_{j=1}^N J_{ij} S(X_{t^{(l)} - \tau(r_i, r_j)}^{j, N}) \right) dt + \lambda(r_i) dW_t^i, \\ t^{(l)} := \sup \{t_l \in \Delta_\delta | t_l \leq t\}, \\ \text{Law of } (x_t)_{t \in [-\tau, 0]} = \bigotimes_{i=1}^N \mu_0(r_i), \end{cases}$$

and by  $Q_0^{N, \delta}$  its Gaussian counterpart.

The synaptic weights are assumed independent and with a law satisfying the Lindenberg-type hypothesis  $(H_J)$  introduced above and that we repeat here in an equivalent manner (see Appendix of [32]):

$$(37) \quad \begin{cases} \exists a, D_0 > 0, \forall N \geq 1, \forall m \leq N, \forall (J_1, \dots, J_m) \in \{\mathcal{L}(J_{ij}(N)), i, j \in \llbracket 1, N \rrbracket\}^m \text{ independent,} \\ \forall (\lambda_1, \dots, \lambda_m) \in [-1, 1]^m, \\ \mathcal{E}_J \left( \exp \frac{aN}{m} (\lambda_1 J_1 + \dots + \lambda_m J_m)^2 \right) \leq D_0. \end{cases}$$

For simplicity of notations, we introduce

$$Y_i := \int_0^T \underbrace{\left( \sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S(x_{t-\tau(r_i, r_j)}^j) \right)}_{\hat{G}_t(r_i)} dW_t(x^i, r_i) - \frac{1}{2} \int_0^T \left( \sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S(x_{t-\tau(r_i, r_j)}^j) \right)^2 dt,$$



$$Y_i^\delta := \int_0^T \underbrace{\left( \sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S(x_{t^{(i)}}^j - \tau(r_i, r_j)) \right)}_{\hat{G}_t^\delta(r_i)} dW_t(x^i, r_i) - \frac{1}{2} \int_0^T \left( \sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S(x_{t^{(i)}}^j - \tau(r_i, r_j)) \right)^2 dt.$$

For  $\epsilon > 0$ , we are interested in the probability  $Q^N(\hat{\mu}_N \in B(Q, \epsilon)^c)$ . In the same spirit as done in (29) we have:

$$\begin{aligned} Q^N(\hat{\mu}_N \in B(Q, \epsilon)^c) &\leq \int_{\{\hat{\mu}_N \notin B(Q, \epsilon)\} \cap A} \frac{dQ^N}{dQ^{N, \delta}} dQ^{N, \delta} + Q^N(A^c) \\ &\leq \underbrace{\left( \int \mathcal{E}_J \left( \mathbb{1}_A \prod_{i=1}^N \exp\{\omega_1(Y_i - Y_i^\delta)\} \exp\{Y_i^\delta\} \right) dP^{\otimes N} \right)^{\frac{1}{\omega_1}}}_{B_A^\delta} Q^{N, \delta}(\{\hat{\mu}_N \in B(Q, \epsilon)^c\} \cap A)^{\frac{1}{\omega_2}} + Q^N(A^c), \\ (38) \quad &\leq \exp\left(-\frac{b_0 N}{p\omega_2}\right) \underbrace{\left( \int_A \left( \frac{dQ^{N, \delta}}{dQ_0^{N, \delta}} \right)^{q-1} dQ^{N, \delta} \right)^{\frac{1}{q\omega_2}}}_{=: Z_N} B_A^\delta \frac{1}{\omega_1} + Q^N(A^c). \end{aligned}$$

using, in the same vein as in (29), a control of the discrete non-Gaussian probability by its discrete Gaussian equivalent, and make use of the exponential decay of the latter.

Hence, proving that the quotients  $\frac{dQ^{N, \delta}}{dQ_0^{N, \delta}}$  and  $\frac{dQ^N}{dQ^{N, \delta}}$  are sufficiently close to 1 on a suitable set  $A$ , the terms in  $Z_N$  and  $B_A^\delta$  would be overridden by the exponential decay. The proof will then be completed by showing that the extra term vanishes.

This is proven in the three following lemmas. The first one controls the term  $Z_N$ , the second one copes with the term  $B_A^\delta$ , whereas the third proves that the extra term vanishes. All this is done for a suitable choice of the set  $A$ .

Let

$$\begin{aligned} a_i^\delta(\mathbf{x}, \mathbf{r}) &:= \mathcal{E}_J \left[ \exp \left( \int_0^T \left( \frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(x_{t^{(i)}}^j - \tau(r_i, r_j)) \right) dW_t(x^i, r_i) \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \int_0^T \left( \frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(x_{t^{(i)}}^j - \tau(r_i, r_j)) \right)^2 dt \right) \right], \\ b_i^\delta(\mathbf{x}, \mathbf{r}) &:= \mathcal{E}_J \left[ \exp \left( \int_0^T \left( \frac{1}{\lambda(r_i)} \sum_{j=1}^N \tilde{J}_{ij} S(x_{t^{(i)}}^j - \tau(r_i, r_j)) \right) dW_t(x^i, r_i) \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \int_0^T \left( \frac{1}{\lambda(r_i)} \sum_{j=1}^N \tilde{J}_{ij} S(x_{t^{(i)}}^j - \tau(r_i, r_j)) \right)^2 dt \right) \right], \end{aligned}$$

so that

$$\frac{dQ^{N, \delta}}{dQ_0^{N, \delta}}(\mathbf{x}, \mathbf{r}) = \prod_{i=1}^N \frac{a_i^\delta(\mathbf{x}, \mathbf{r})}{b_i^\delta(\mathbf{x}, \mathbf{r})}.$$

**Lemma 21.** *There exists a set  $A_{N,\delta}^2 \in \mathcal{B}((\mathcal{C} \times D)^N)$  with  $P^{\otimes N}(A_{N,\delta}^2) = 1$ , on which the  $(a_i^\delta)$  and  $(b_i^\delta)$  satisfy the following properties:*

$$(H1) \quad \exists A, B > 0, \forall N, \frac{1}{\delta} \in \mathbb{N}, i \leq N, \quad a_i^\delta(\mathbf{x}, \mathbf{r}) \geq A \exp \left( -\sqrt{\delta} B \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_i, r_i)| \right)$$

$$(H2) \quad \exists \lambda < 1, C > 0, \forall N, \frac{1}{\delta} \in \mathbb{N}, i \leq N, \quad a_i^\delta(\mathbf{x}, \mathbf{r}) \leq C \exp \left( \frac{\lambda}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right)$$

(H3)

$\forall \eta > 0, \exists \alpha \in [0, 1], \forall N \geq 1, \forall \frac{1}{\delta} \in \mathbb{N}, k \leq N$ , if  $\frac{k}{N} \leq \alpha$  then,  $\forall s$ , injection from

$\{1, \dots, k\}$  into  $\{1, \dots, N\}$ ,  $\forall i \notin \text{Im}(s) = \{s(1), \dots, s(k)\}$ ,  $\exists \tilde{a}_i^\delta$  depending only on  $j \notin \text{Im}(s)$ , such that,

$$\sup \left( \frac{a_i^\delta(\mathbf{x}, \mathbf{r})}{\tilde{a}_i^\delta(\mathbf{x}, \mathbf{r})}, \frac{\tilde{a}_i^\delta(\mathbf{x}, \mathbf{r})}{a_i^\delta(\mathbf{x}, \mathbf{r})} \right) \leq (1 + \eta) \exp \left( \frac{\eta}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right)$$

$$(H4) \quad \exists D > 0, \forall \beta > 0, \exists N_0, \forall \frac{1}{\delta} \in \mathbb{N}, \forall N \geq N_0, \forall i \in \{1, \dots, N\},$$

$$\frac{a_i^\delta(\mathbf{x}, \mathbf{r})}{b_i^\delta(\mathbf{x}, \mathbf{r})} \leq 1 + \beta \exp \left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left( B_{t_l}^2(x_i, r_i) + D\sqrt{\delta} |B_{t_l}(x_i, r_i)| \right) \right\},$$

where  $B_{t_l}(x, r) = \left( \frac{W_{t_{l+1}}(x, r) - W_{t_l}(x, r)}{\sqrt{\delta T}} \right)$ .

Moreover, on  $A_{N,\delta}^0$ , choosing  $\eta_0 > 0$  such that  $\alpha + 2\eta_0 < 1$ ,  $\eta \leq \eta_0$ ,  $\alpha \leq \eta$ , then  $\exists C_1, C_2 > 0, \xi < 1$  such that

$$Z_N \leq (1 + \beta^{q-1} C_{\alpha,\delta})^N (1 + \eta)^{2N} \left( \exp \left\{ N \left( \frac{\sqrt{\eta}}{\delta} + B \frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}} \right) \right\} + C_2^{\frac{N}{\delta}} \exp \left\{ N \frac{\xi - 1}{4\delta\sqrt{\eta}} \right\} \right),$$

where  $C_{\alpha,\delta} := \max \left( C_1^{\frac{1}{\alpha\delta}}, \frac{C}{A} C_1^{\frac{1}{\delta}}, 1 \right)$ .

*Remark 8.* Sums of squares of centered standard Gaussian appear in an exponential for the upperbound of (H2), (H3) and (H4). There are exactly as many Gaussians as points in the partition  $\Delta_\delta$ , so that in the continuous limit, these terms will diverge.

We will demonstrate that all the hypotheses of lemma 21 are valid in the case of the randomly connected network. Their proof are postponed in Appendix D.

**Lemma 22.** *For  $\frac{T}{\lambda_*^2} < a$  and any conjugate exponents  $(\kappa_1, \kappa_2)$ , with  $\kappa_1 - 1, \omega_1 - 1$  small enough, exists a constant  $C_T$  independent of  $N$  and  $\delta$ , and a set  $A_{N,\delta}^1 \in \mathcal{B}((\mathcal{C} \times D)^N)$  such that, if  $\kappa_2 = O(\delta^{-\frac{1}{4}})$  then:*

$$\exists \delta_0 > 0, \forall \delta < \delta_0, \forall N, \quad B_{A_{N,\delta}^1}^\delta \leq \exp \left\{ C_T N (\kappa_1 - 1) \right\}.$$

Restricting the integral on the set  $A_{N,\delta} := A_{N,\delta}^1 \cap A_{N,\delta}^2$  will allow obtaining a proper control on both the  $a_i^\delta$ , the  $b_i^\delta$  and the term  $B_{A_{N,\delta}^1}^\delta$ . It also makes  $Q^N(A_{N,\delta}^c)$  appear in (38). We must justify that this quantity goes to zero as  $N$  grows to infinity. This is the purpose of the following lemma, whose proof is postponed to the end of the section:

**Lemma 23.** *For  $T < \lambda_*^2 a$ , exists a constant  $\delta_0 > 0$  such that  $\forall \delta < \delta_0$ ,  $Q^N(A_{N,\delta}^c)$  decreases exponentially fast to zero as  $N$  goes to infinity.*

We now state and prove the main result of the section:

**Theorem 24.** *Let  $(J_{ij})$  satisfy condition  $(H_J)$  and  $\delta$  be small enough. Then, under a short time hypothesis, the empirical measure undergoes, under  $Q^N$ , a convergence in law toward  $Q$ . Namely, if  $\frac{T}{\lambda_*^2} < a$ , then,*

$$\exists b > 0, \exists N_0 \in \mathbb{N}, \forall N \geq N_0, \quad Q^N(\hat{\mu}_N \in B(Q, \epsilon)^c) \leq \exp(-bN).$$

*Proof of Theorem 24.* We start by choosing in a specific order the parameters appearing in the previous estimates. For any  $\delta > 0$  (small enough), let  $b_0 > 0$  such that

$$(39) \quad Q^{N,\delta}(A_{N,\delta} \cap \{\hat{\mu}_N \in B(Q, \epsilon)^c\}) \leq \exp\left(-\frac{b_0 N}{p}\right) Z_N^{\frac{1}{q}}.$$

Let  $\eta = \delta^3$ , and  $\alpha_\delta \leq \eta$  as in hypothesis (H3). Remark that it is valid for any pair  $(\frac{1}{\delta}, N)$  with  $N \in \mathbb{N}$ . Moreover, let  $0 < \gamma < \frac{1}{4}$  and fix  $\omega_1 = 1 + \delta^{\frac{\gamma}{2}}$ ,  $\kappa_1 = 1 + \delta^\gamma$  so that  $\kappa_2 \sim \frac{1}{\delta^\gamma}$ . Then, for  $\delta$  small enough and under the short time condition  $\frac{T}{\lambda_*^2} < a$ , the conditions of Lemmas 22 and 23, as well as  $\eta < \eta_0$  will be satisfied, with

$$B^\delta \leq \exp\{C_T N \delta^\gamma\}.$$

Furthermore, inequality (38) gives

$$Q^N(\hat{\mu}_N \in B(Q, \epsilon)^c) \leq \exp\left(-\frac{b_0 N(\omega_1 - 1)}{p\omega_1}\right) \exp\{C_T N \delta^\gamma\} Z_N^{\frac{\omega_1 - 1}{q\omega_1}} + Q^N(A_{N,\delta}^c).$$

Lemma 21 ensures that for  $\delta$  small enough, one has  $\frac{1-\xi}{4\sqrt{\eta}} > \log(C_2)$ , so that

$$Z_N \leq 2 \exp\{N \log(1 + \beta^{q-1} C_{\alpha,\delta})\} \exp\{2N((1+B)\delta^{\frac{1}{4}} + \log(1 + \delta^3))\}.$$

Moreover, for  $\delta$  small enough,  $C_{\alpha,\delta} = C_1^{\frac{1}{\alpha\delta}}$ , so that choosing  $\beta = C_{\alpha,\delta}^{\frac{-2}{q-1}}$ ,  $Z_N$  is at most of order  $\exp\{C_T \delta^{\frac{1}{4}} N\}$ . As  $\gamma < \frac{1}{4}$ , one can see that for  $\delta$  small enough, exists  $N_0$  such that  $\forall N \geq N_0$

$$Q^N(\hat{\mu}_N \in B(Q, \epsilon)^c) \leq \exp(-CN \delta^{\frac{\gamma}{2}}) + Q^N(A_{N,\delta}^c).$$

□

We now prove Lemmas 22, 23 and then 21.

*Proof of Lemma 22.* We use Hölder inequality with conjugate exponents  $(\kappa_1, \kappa_2)$ , to split  $B_A^\delta$  in two terms:

$$B_A^\delta \leq \underbrace{\left\{ \int \prod_{i=1}^N \exp\{\kappa_1 Y_i\} dP^{\otimes N} \right\}^{\frac{1}{\kappa_1}}}_{B_1^\delta} \underbrace{\left\{ \int_A \prod_{i=1}^N \mathcal{E}_J \left( \exp\{(\omega_1 - 1)\kappa_2(Y_i - Y_i^\delta)\} \right) dP^{\otimes N} \right\}^{\frac{1}{\kappa_2}}}_{B_2^\delta}.$$

To control each of these terms, we will mainly rely on martingale property, as well as on the hypothesis  $(H_J)$ . For the control of the terms  $B_1^\delta$ , the idea is to chose  $\kappa_1$  sufficiently close to 1 so that the expectation of the power of the martingale  $\prod_{i=1}^N \exp\{Y_i\}$  will be lamost equal to 1. The smallness of the second term will be a consequence of the Hölder continuity of solutions under  $P$ . In detail, we have:

$$\begin{aligned} B_1^\delta &= \mathcal{E}_J \left( \int \prod_{i=1}^N \exp \left\{ \kappa_1 \int_0^T \hat{G}_t^\delta(r_i) dW_t(x^i, r_i) - \frac{\kappa_1}{2} \int_0^T \hat{G}_t^\delta(r_i)^2 dt \right\} dP^{\otimes N} \right) \\ &\stackrel{\text{Hölder}}{\leq} \mathcal{E}_J \left( \int \prod_{i=1}^N \exp \left\{ \kappa_1^2 \int_0^T \hat{G}_t^\delta(r_i) dW_t(x^i, r_i) - \frac{\kappa_1^4}{2} \int_0^T \hat{G}_t^\delta(r_i)^2 dt \right\} dP^{\otimes N} \right)^{\frac{1}{\kappa_1}} \\ &\quad \times \mathcal{E}_J \left( \int \prod_{i=1}^N \exp \left\{ \frac{\kappa_1}{\kappa_1 - 1} \frac{\kappa_1^3 - \kappa_1}{2\lambda_*^2} \int_0^T \left( \sum_{j=1}^N J_{ij} S(x_{t^{(i)} - \tau(r_i, r_j)}^j) \right)^2 dt \right\} dP^{\otimes N} \right)^{\frac{\kappa_1 - 1}{\kappa_1}}. \end{aligned}$$

The first term of the right-hand side is equal to one by martingale property, so that using Fubini theorem and Jensen inequality yields

$$B_1^\delta \leq \left\{ \int \prod_{i=1}^N \int_0^T \mathcal{E}_J \left( \exp \left\{ \frac{\kappa_1^2(\kappa_1 + 1)T}{2\lambda_*^2} \left( \sum_{j=1}^N J_{ij} S(x_{t^{(i)} - \tau(r_i, r_j)}^j) \right)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N} \right\}^{\frac{\kappa_1 - 1}{\kappa_1}}.$$

Moreover, using hypothesis  $(H_J)$  and the inequality  $\frac{T}{\lambda_*^2} < a$ , one can choose  $\kappa_1 - 1$  small enough so to obtain

$$(40) \quad (B_1^\delta)^{\frac{1}{\kappa_1}} \leq \exp \{ C_T(\kappa_1 - 1)N \}.$$

We now deal with the second term:

$$\begin{aligned} B_2^\delta &= \mathcal{E}_J \left( \int_A \prod_{i=1}^N \exp \left\{ (\omega_1 - 1)\kappa_2 \int_0^T \left( \hat{G}_t(r_i) - \hat{G}_t^\delta(r_i) \right) dW_t(x^i, r_i) \right. \right. \\ &\quad \left. \left. - \frac{(\omega_1 - 1)\kappa_2}{2} \int_0^T \hat{G}_t(r_i)^2 - \hat{G}_t^\delta(r_i)^2 dt \right\} dP^{\otimes N} \right) \\ &\stackrel{\text{C.S.}}{\leq} \mathcal{E}_J \left( \int \prod_{i=1}^N \exp \left\{ 2(\omega_1 - 1)\kappa_2 \int_0^T \left( \hat{G}_t(r_i) - \hat{G}_t^\delta(r_i) \right) dW_t(x^i, r_i) \right. \right. \\ &\quad \left. \left. - 4(\omega_1 - 1)^2 \kappa_2^2 \int_0^T \left( \hat{G}_t(r_i) - \hat{G}_t^\delta(r_i) \right)^2 dt \right\} dP^{\otimes N} \right)^{\frac{1}{2}} \\ &\quad \times \mathcal{E}_J \left( \int_A \prod_{i=1}^N \exp \left\{ 4(\omega_1 - 1)^2 \kappa_2^2 \int_0^T \left( \hat{G}_t(r_i) - \hat{G}_t^\delta(r_i) \right)^2 dt \right. \right. \\ &\quad \left. \left. - (\omega_1 - 1)\kappa_2 \int_0^T \hat{G}_t(r_i)^2 - \hat{G}_t^\delta(r_i)^2 dt \right\} dP^{\otimes N} \right)^{\frac{1}{2}}. \end{aligned}$$

As previously, the first term of the right-hand side is equal to 1 by the martingale property. Moreover,

$$-(\hat{G}^2 - (\hat{G}^\delta)^2) = (\hat{G}^\delta - \hat{G})(\hat{G} + \hat{G}^\delta) \leq \frac{\kappa_2}{2}(\hat{G}^\delta - \hat{G})^2 + \frac{1}{2\kappa_2}(\hat{G}^\delta + \hat{G})^2.$$

Hypothesis  $(H_J)$  allow to control the first of these terms:

$$\begin{aligned}
(\hat{G}^\delta - \hat{G})^2 &\leq \frac{K_S^2 \delta^{\frac{1}{2}}}{\lambda_*^2} \left( \sum_{j=1}^N J_{ij} \underbrace{\frac{S(x_{t-\tau(r_i, r_j)}^j) - S(x_{t^{(l)}-\tau(r_i, r_j)}^j)}{K_S \delta^\beta}}_{\lambda_{ij}(t)} \right)^2. \\
B_2^\delta &\stackrel{\text{Fubini}}{\leq} \left\{ \int_A \prod_{i=1}^N \mathcal{E}_J \left( \exp \left\{ \frac{K_S^2 (\omega_1 - 1) \kappa_2^2 \delta^{2\beta} (8(\omega_1 - 1) + 1)}{2\lambda_*^2} \int_0^T \left( \sum_{j=1}^N J_{ij} \lambda_{ij}(t) \right)^2 dt \right. \right. \right. \\
&\quad \left. \left. + \frac{\omega_1 - 1}{2} \int_0^T (\hat{G}_t(r_i) + \hat{G}_t^\delta(r_i))^2 dt \right\} \right) dP^{\otimes N} \right\}^{\frac{1}{2}} \\
&\stackrel{\text{C.S.}}{\leq} \left\{ \int_A \prod_{i=1}^N \mathcal{E}_J \left( \int_0^T \exp \left\{ C_T (\omega_1 - 1) \kappa_2^2 \delta^{2\beta} \left( \sum_{j=1}^N J_{ij} \lambda_{ij}(t) \right)^2 \right\} \frac{dt}{T} \right) dP^{\otimes N} \right\}^{\frac{1}{4}} \\
&\times \left\{ \int \prod_{i=1}^N \mathcal{E}_J \left( \int_0^T \exp \left\{ C_T (\omega_1 - 1) \left( \sum_{j=1}^N J_{ij} \frac{S(x_{t-\tau(r_i, r_j)}^j) + S(x_{t^{(l)}-\tau(r_i, r_j)}^j)}{2} \right)^2 \right\} \frac{dt}{T} \right) dP^{\otimes N} \right\}^{\frac{1}{4}}.
\end{aligned}$$

One sees that, for  $\omega_1 - 1$  small enough, the second term in the right-hand side can be handled using again Fubini theorem and hypothesis  $(H_J)$ . We split the other term into two parts: one in which we keep only the  $\lambda_{ij}$  that behave nicely, so that we can rely on hypothesis  $(H_J)$ , and the other one in which only pathological  $\lambda_{ij}$  appear: these may bring large contributions, but they appear infrequently. Moreover, even for such  $\lambda_{ij}$ ,  $K_S \delta^{\frac{1}{4}} |\lambda_{ij}| \leq 2$ . In a more formal way, remember the definitions of  $E_{N,\delta}^j$ ,  $c_{N,\delta}$  and  $A_{N,\delta}^1$  introduced in the proof of Theorem 13. In particular, on the event  $E_{N,\delta}^j$ , every quantity  $|\lambda_{ij}(t)|$  is smaller than 1 for  $\delta$  small enough. Then

$$\begin{aligned}
&\left\{ \int_{A_{N,\delta}^1} \prod_{i=1}^N \int_0^T \mathcal{E}_J \left( \exp \left\{ C_T (\omega_1 - 1) \kappa_2^2 \delta^{2\beta} \left( \sum_{j=1}^N J_{ij} \lambda_{ij}(t) \right)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N} \right\}^{\frac{1}{4}} \\
&\stackrel{\text{C.S.}}{\leq} \left\{ \int \prod_{i=1}^N \int_0^T \mathcal{E}_J \left( \exp \left\{ C_T (\omega_1 - 1) \kappa_2^2 \delta^{2\beta} \left( \sum_{j=1}^N \mathbf{1}_{E_{N,\delta}^j} J_{ij} \lambda_{ij}(t) \right)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N} \right\}^{\frac{1}{8}} \\
&\times \left\{ \int_{A_{N,\delta}^1} \prod_{i=1}^N \int_0^T \mathcal{E}_J \left( \int_0^T \exp \left\{ C_T (\omega_1 - 1) \kappa_2^2 \left( \sum_{j=1}^N \mathbf{1}_{(E_{N,\delta}^j)^c} J_{ij} \frac{K_S \delta^{\frac{1}{4}} \lambda_{ij}(t)}{2} \right)^2 \right\} \frac{dt}{T} dP^{\otimes N} \right) \right\}^{\frac{1}{8}}.
\end{aligned}$$

Remark that, considered under  $P^{\otimes N}$ , the  $\lambda_{ij}$  are independent of the matrix  $J$ . Then, for  $\kappa_2 = O(\delta^{-\beta})$ , hypothesis  $(H_J)$  yields

$$\left\{ \int \prod_{i=1}^N \int_0^T \mathcal{E}_J \left( \exp \left\{ C_T (\omega_1 - 1) \kappa_2^2 \delta^{2\beta} \left( \sum_{j=1}^N \mathbf{1}_{E_{N,\delta}^j} J_{ij} \lambda_{ij}(t) \right)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N} \right\}^{\frac{1}{8}} \leq \exp \left\{ C_T (\omega_1 - 1) \kappa_2^2 \delta^{2\beta} N \right\}.$$

Now for the other term, as  $c_{N,\delta} \leq \delta^{2\beta} N$  on  $A_{N,\delta}^1$

$$\left\{ \int_{A_{N,\delta}^1} \prod_{i=1}^N \mathcal{E}_J \left( \int_0^T \exp \left\{ C_T(\omega_1 - 1) \kappa_2^2 \left( \sum_{j=1}^N \mathbf{1}_{(E_{N,\delta}^j)^c} J_{ij} \frac{K_S \delta^\beta \lambda_{ij}(t)}{2} \right)^2 \right\} \frac{dt}{T} \right) dP^{\otimes N} \right\} \leq$$

$$\left\{ \int_{A_{N,\delta}^1} \prod_{i=1}^N \int_0^T \mathcal{E}_J \left( \exp \left\{ C_T(\omega_1 - 1) \kappa_2^2 \delta^{2\beta} \frac{N}{c_{N,\delta}} \left( \sum_{j=1}^N \mathbf{1}_{(E_{N,\delta}^j)^c} J_{ij} \frac{K_S \delta^\beta \lambda_{ij}(t)}{2} \right)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N} \right\}.$$

Moreover,  $\frac{K_S \delta^\beta |\lambda_{ij}(t)|}{2} \leq 1$ , so that for  $\kappa_2 = O(\delta^{-\beta})$ , we are also in the scope of hypothesis  $(H_J)$ .

To summarize, as soon as  $\omega_1 - 1$  is small enough and for  $\kappa_2 = O(\delta^{-\beta})$  one can use Fubini Theorem and hypothesis  $(H_J)$  to obtain

$$(41) \quad (B_2^\delta)^{\frac{1}{\kappa_2}} \leq \exp \left\{ C_T \left( (\omega_1 - 1) \kappa_2 \delta^{2\beta} + \frac{(\omega_1 - 1)}{\kappa_2} \right) N \right\} \leq \exp \left\{ \frac{C_T(\omega_1 - 1)N}{\kappa_2} \right\}.$$

Hence, using inequalities (40) and (41) with  $\kappa_1 - 1$ ,  $\omega_1 - 1$  small enough,  $\kappa_2 = O(\delta^{-\beta})$ , and under a short time hypothesis  $\frac{T}{\lambda_*^2} < a$ , there exists a constant  $C_T$  independent of  $N$  and  $\delta$  such that

$$B_{A_{N,\delta}^1}^\delta \leq \exp \left\{ C_T \left( (\kappa_1 - 1) + \frac{\omega_1 - 1}{\kappa_2} \right) N \right\} \leq \exp \left\{ C_T(\kappa_1 - 1)N \right\}.$$

□

Let us now take care of  $Z_N$  appearing in the right-hand side of (38).

*Proof of Lemma 21.* Assumption (H1)- (H4) are demonstrated in Appendix D. We prove here the inequality involving  $Z_N$ . Let  $\lambda$  be the constant of condition (H2). Then, we chose in (38) conjugate exponents  $(p, q)$  satisfying  $q \in ]1, \frac{3}{2}[$  and  $\lambda + (q - 1) < 1$ . Then,

$$Z_N \leq \int_{A_{N,\delta}^2} \prod_{i=1}^N \left( 1 + \frac{a_i^\delta - b_i^\delta}{b_i^\delta} \right)^{q-1} dQ^{N,\delta}(\mathbf{x}, \mathbf{r}).$$

Property (H4) implies that  $\left| \frac{a_i^\delta - b_i^\delta}{b_i^\delta} \right| \leq \beta \exp \left\{ \frac{1}{2} \sum_{l=0}^{\frac{1}{\delta}} \left( B_{t_l}^2(x_i, r_i) + D\sqrt{\delta} |B_{t_l}(x_i, r_i)| \right) \right\}$ . Moreover, as  $(x + y)^{q-1} \leq x^{q-1} + y^{q-1}$  for any  $x, y > 0$ , one has

$$(42) \quad Z_N \leq 1 + \sum_{k=1}^N \frac{\beta^{k(q-1)}}{k!} \sum_{s \in \mathcal{I}_N^k} O_{s,k},$$

where  $\mathcal{I}_N^k$  is the set of injective application from  $\llbracket 1, k \rrbracket$  to  $\llbracket 1, N \rrbracket$ , and

$$O_{s,k} = \int_{A_{N,\delta}^2} \prod_{i=1}^k \exp \left\{ \frac{q-1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left( B_{t_l}^2(x_{s(i)}, r_{s(i)}) + \sqrt{\delta} D |B_{t_l}(x_{s(i)}, r_{s(i)})| \right) \right\} dQ^{N,\delta}(\mathbf{x}, \mathbf{r}).$$

Let  $\eta > 0$ , and  $\alpha \leq \eta$  be as in (H3). Then, if  $\frac{k}{N} > \alpha$  we can apply (H2) to obtain:

$$\begin{aligned}
O_{s,k} &\leq O_{s,N} \leq C^N \int_{A_{N,\delta}^2} \prod_{i=1}^N \left( \exp \left\{ \frac{\lambda + (q-1)}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) + \sqrt{\delta} D \frac{q-1}{2} \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_i, r_i)| \right\} \right) dP^{\otimes N}(\mathbf{x}, \mathbf{r}) \\
&\leq \left( C \prod_{l=1}^{\frac{1}{\delta}} \int_{A_{N,\delta}^2} \exp \left\{ \frac{\lambda + (q-1)}{2} B_{t_l}^2(x, r) + \sqrt{\delta} D \frac{q-1}{2} |B_{t_l}(x, r)| \right\} dP(x, r) \right)^N \leq C_1^{\frac{N}{\delta}}
\end{aligned}$$

so that

$$O_{s,k} \leq (C_1^{\frac{1}{\alpha\delta}})^k.$$

Suppose now that  $\frac{k}{N} \leq \alpha$ . We then use property (H3) for  $a_i^\delta$ , with  $i \notin \{s(1), \dots, s(k)\}$ , and property (H2) for the other  $i$ , and obtain by independence

$$O_{s,k} \leq (1 + \eta)^N F_N G_N$$

where

$$F_N = C^k \int_{A_{N,\delta}^2} \prod_{i=1}^k \left( \exp \left\{ \frac{\lambda + (q-1)}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_{s(i)}, r_{s(i)}) + \sqrt{\delta} D \frac{q-1}{2} \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_{s(i)}, r_{s(i)})| \right\} \right) dP^{\otimes N}(\mathbf{x}, \mathbf{r}),$$

and

$$G_N = \int_{A_{N,\delta}^2} \prod_{i \notin Im(s)} \tilde{a}_i^\delta \exp \left\{ \frac{\eta}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right\} dP^{\otimes N}.$$

As previously,  $F_N \leq C_1^{\frac{k}{\delta}}$ . Moreover, using (H3) and (H1) to recover every  $a_i^\delta$ , we obtain

$$G_N \leq \frac{(1 + \eta)^N}{A^k} \int_{A_{N,\delta}^2} \left( \prod_{i=1}^N \exp \left\{ \eta \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right\} \right) \left( \prod_{i=1}^k \exp \left\{ B \sqrt{\delta} \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_i, r_i)| \right\} \right) dQ^{N,\delta}.$$

Let now

$$I_N = \int_{A_{N,\delta}^2} \exp \left\{ \eta \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) + \sqrt{\delta} B \sum_{i=1}^k \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_{s(i)}, r_{s(i)})| \right\} \mathbf{1}_{\{\sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \leq \frac{N}{\sqrt{\eta\delta}}\}} dQ^{N,\delta},$$

and

$$J_N = \int_{A_{N,\delta}^2} \exp \left\{ \eta \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) + \sqrt{\delta} B \sum_{i=1}^k \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_{s(i)}, r_{s(i)})| \right\} \mathbf{1}_{\{\sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) > \frac{N}{\sqrt{\eta\delta}}\}} dQ^{N,\delta},$$

so that

$$G_N \leq \frac{(1 + \eta)^N}{A^k} (I_N + J_N).$$

As  $\alpha \leq \eta$ , and

$$\sum_{i=1}^k \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_{s(i)}, r_{s(i)})| \stackrel{C.S.}{\leq} \left( \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right)^{\frac{1}{2}} \cdot \sqrt{\frac{k}{\delta}},$$

we finds

$$I_N \leq \exp \left\{ N \left( \frac{\sqrt{\eta}}{\delta} + B \frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}} \right) \right\}.$$

Moreover, for  $\eta \leq \eta_0$ , with  $\lambda + 2\eta_0 < 1$ , using (H2) we have

$$J_N \leq \int_{A_{N,\delta}^2} \exp \left\{ \frac{\lambda + 2\eta}{2} \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) + B\sqrt{\delta} \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_i, r_i)| \right\} \mathbf{1}_{\{\sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) > \frac{N}{\sqrt{\eta}\delta}\}} dP^{\otimes N}.$$

Under  $P^{\otimes N}$  the  $(B_{t_l}^2(x_i, r_i))_{i,l}$  are independent centered standard Gaussian variables. Hence, writting down their density, we see that exists  $\xi < 1$  such that

$$\begin{aligned} J_N &\leq \int_{A_{N,\delta}^2} \exp \left\{ \frac{\xi - 1}{2} \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} u_{i,l}^2 + B\sqrt{\delta} \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} |u_{i,l}| \right\} \mathbf{1}_{\{\sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} u_{i,l}^2 > \frac{N}{\sqrt{\eta}\delta}\}} \frac{du_{1,1} \dots du_{N,\frac{1}{\delta}}}{\sqrt{(2\pi)^{\frac{N}{\delta}}}}, \\ J_N &\leq \exp \left\{ N \frac{\xi - 1}{4\delta\sqrt{\eta}} \right\} \int_{A_{N,\delta}^2} \exp \left\{ \frac{\xi - 1}{4} \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} u_{i,l}^2 + B\sqrt{\delta} \sum_{i=1}^N \sum_{l=1}^{\frac{1}{\delta}} |u_{i,l}| \right\} \frac{du_{1,1} \dots du_{N,\frac{1}{\delta}}}{\sqrt{(2\pi)^{\frac{N}{\delta}}}}, \\ J_N &\leq C_2^{\frac{N}{\delta}} \exp \left\{ N \frac{\xi - 1}{4\delta\sqrt{\eta}} \right\}. \end{aligned}$$

Therefore, letting  $C_{\alpha,\delta} = \max(C_1^{\frac{1}{\alpha\delta}}, \frac{C}{A}C_1^{\frac{1}{\delta}}, 1)$ , we obtain  $\forall k \in \{1, \dots, N\}$

$$O_{s,k} \leq (1 + \eta)^{2N} C_{\alpha,\delta}^k \left( \exp \left\{ N \left( \frac{\sqrt{\eta}}{\delta} + B \frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}} \right) \right\} + C_2^{\frac{N}{\delta}} \exp \left\{ N \frac{\xi - 1}{4\delta\sqrt{\eta}} \right\} \right)$$

and injecting this upperbound in (42) we get

$$Z_N \leq (1 + \beta^{q-1} C_{\alpha,\delta})^N (1 + \eta)^{2N} \left( \exp \left\{ N \left( \frac{\sqrt{\eta}}{\delta} + B \frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}} \right) \right\} + C_2^{\frac{N}{\delta}} \exp \left\{ N \frac{\xi - 1}{4\delta\sqrt{\eta}} \right\} \right).$$

□

*Proof of Lemma 23.* As  $P^{\otimes N}(A_{N,\delta}^2) = 1$  as soon as  $N \geq \frac{1}{\delta}$  (see (47)), it suffices to show the result on  $A_{N,\delta}^1$ . Hölder inequality yields

$$\int \mathbf{1}_{(A_{N,\delta}^1)^c} \frac{dQ^N}{dP^{\otimes N}} dP^{\otimes N} \stackrel{\text{C.S.}}{\leq} \left( \int \exp\{\kappa_1 N \Gamma(\hat{\mu}_N)\} dP^{\otimes N} \right)^{\frac{1}{\kappa_1}} P^{\otimes N}((A_{N,\delta}^1)^c)^{\frac{1}{\kappa_2}}.$$

As done in the proof of Lemma 22, we obtain under a short-time hypothesis that

$$\left( \int \exp\{\kappa_1 N \Gamma(\hat{\mu}_N)\} dP^{\otimes N} \right)^{\frac{1}{\kappa_1}} \leq \exp\{C_T(\kappa_1 - 1)N\},$$

where  $\tilde{C}_T$  is independent of  $N$  and  $\delta$ . Furthermore,

$$P^{\otimes N}((A_{N,\delta}^1)^c) = P^{\otimes N}(c_{N,\delta} > \delta^{2\beta} N) = P^{\otimes N} \left( \frac{\sum_{i=1}^N \mathbf{1}_{(E_{N,\delta}^i)^c} - P((E_{N,\delta}^1)^c)}{N} > \delta^{2\beta} - P((E_{N,\delta}^1)^c) \right).$$

Let us show that  $\forall m \in \mathbb{N}^*, \exists C_{T,m} > 0$  such that  $P((E_{N,\delta}^1)^c) < C_{T,m} \delta^{m(\frac{1}{2}-\beta)-1}$ ,

so that for  $m$  big enough and  $\delta$  small enough,  $P((E_{N,\delta}^1)^c) < \delta^{2\beta}$ .

Remember that the semi-martingale decomposition of  $x$  under  $P_r$

$$x_t - x_s = \int_s^t f(r, u, x_u) du + \lambda(r)(W_t(x, r) - W_s(x, r))$$



so that using the Lipschitz continuity of  $S$ , one has

$$P((E_{N,\delta}^1)^c) \leq P\left(\sup_{s \in [-\tau, T-\delta], t \in [s, s+\delta]} \left| \int_s^t f(r_1, u, x_u^1) du \right| > \frac{\delta^\beta}{2}\right) \\ + P\left(\sup_{s \in [-\tau, T-\delta], t \in [s, s+\delta]} |W_t(x^1, r_1) - W_s(x^1, r_1)| > \frac{\delta^\beta}{2\lambda^*}\right).$$

In fact suppose we are on  $(E_{N,\delta}^1)^c$ , that is there exist  $s \in [-\tau, T-\delta]$  and  $t \in [s, s+\delta]$  such that  $|x_t^1 - x_s^1| > \delta^\beta$ . Then, one scenario is that  $s^{(l)} = t^{(l)}$  or  $t = s^{(l)} + \delta$ , so that  $s$  and  $t$  are in the same interval  $[s^{(l)}, s^{(l)} + \delta]$ , while the other possibility is that they belong to two different consecutive such intervals. Then, by triangular inequality

$$P\left(\sup_{s \in [-\tau, T-\delta], t \in [s, s+\delta]} |W_t^1 - W_s^1| > \frac{\delta^\beta}{2\lambda^*}\right) \leq P\left(\sup_{t \in [-\tau, T]} \max(|W_t^1 - W_{t^{(l)}}^1|, |W_t^1 - W_{t^{(l)}+\delta}^1|) > \frac{\delta^\beta}{4\lambda^*}\right) \\ \stackrel{Markov}{\leq} \frac{\mathbb{E}\left[\sup_{t \in [-\tau, T]} \max(|W_t^1 - W_{t^{(l)}}^1|^m, |W_t^1 - W_{t^{(l)}+\delta}^1|^m)\right]}{(4\lambda^*)^m \delta^{m\beta}} \\ \leq (4\lambda^*)^m \frac{\sum_{l=-\tau}^{T/\delta} \mathbb{E}\left[\sup_{t \in [t^{(l)}, t^{(l)}+\delta]} |W_t^1 - W_{t^{(l)}}^1|^m\right] + \mathbb{E}\left[\sup_{t \in [t^{(l)}-\delta, t^{(l)}]} |W_t^1 - W_{t^{(l)}}^1|^m\right]}{\delta^{m\beta}} \\ \leq \frac{2(T+\tau)}{\delta^{m\beta+1}} (4\lambda^*)^m \mathbb{E}\left[\sup_{t \in [0, \delta]} |W_t^1|^m\right] \stackrel{\text{BDG}}{\leq} C_{T,m} \delta^{m(\frac{1}{2}-\beta)-1},$$

where  $C_{T,m}$  is a constant independent of  $\delta$  and  $N$ . Moreover, Markov inequality gives:

$$P\left(\sup_{s \in [-\tau, T-\delta], t \in [s, s+\delta]} \left| \int_s^t f(r_1, u, x_u^1) du \right| > \frac{\delta^\beta}{2}\right) \leq 2^m \delta^{m(1-\beta)} \mathbb{E}\left[\sup_{t \in [-\tau, T]} |f(r_1, t, X_t^{1,N})|^m\right].$$

Furthermore, it is proven in the Appendix A that  $\sup_{t \in [-\tau, T]} |f(r_1, t, X_t^{1,N})|$  admits exponential moments under  $\mathbb{P}$ . Let  $\sigma_{N,\delta}^2 := \text{Var}(\mathbf{1}_{(E_{N,\delta}^1)^c}) = P^{\otimes N}((E_{N,\delta}^1)^c)(1 - P^{\otimes N}((E_{N,\delta}^1)^c)) \leq C_{T,m} \delta^{m(\frac{1}{2}-\beta)-1}$ . Then, for  $\delta$  small enough,

$$P^{\otimes N}((A_{N,\delta}^1)^c) \leq P^{\otimes N}\left(\frac{1}{2\sigma_{N,\delta}\sqrt{N}} \left| \sum_{i=1}^N \mathbf{1}_{(E_{N,\delta}^1)^c} - P((E_{N,\delta}^1)^c) \right| > \frac{1}{4\sigma_{N,\delta}} \sqrt{N} \delta^{2\beta}\right) \\ \leq \exp\{-\tilde{C}_{T,m} N \delta^{2+4\beta-2m(\frac{1}{2}-\beta)}\} \mathbb{E}\left[\exp\left\{\frac{1}{4} \mathcal{N}(0,1)^2\right\}\right].$$

Hence

$$\int \mathbf{1}_{(A_{N,\delta}^1)^c} \frac{dQ^N}{dP^{\otimes N}} dP^{\otimes N} \leq \exp\left\{-\tilde{C}_{T,m} \left(\delta^{2+4\beta-2m(\frac{1}{2}-\beta)} + 1\right) (\kappa_1 - 1) N\right\}.$$

For  $\delta$  small enough, and  $m$  big enough, this term goes to zero exponentially fast with  $N \rightarrow \infty$ .  $\square$

## APPENDIX A. A PRIORI ESTIMATES FOR SINGLE NEURONS

**Lemma 25.**  $\forall r \in D, \sup_{t \in [-\tau, T]} |f(r, t, x_t)|^2$  admits exponential moments under  $P_r$ .

*Proof.* Fix  $r \in D$ , and suppose first that  $t \in [0, T]$ . Using the Lipschitz continuity of  $f(r, \cdot, \cdot)$ , we have

$$\begin{aligned} |f(r, t, x_t)| &\leq |f(r, 0, x_0)| + |f(r, t, x_t) - f(r, 0, x_0)| \\ &\leq |f(r, 0, x_0)| + K_f T + K_f \int_0^t |f(r, u, x_u)| du + K_f \lambda^* |W_t(x, r)| \\ &\leq |f(0, 0, 0)| + K_f(|x_0| + T + d_D) + K_f \int_0^t |f(r, u, x_u)| du + K_f \lambda^* |W_t(x, r)| \end{aligned}$$

so that by Gronwall's lemma:

$$(43) \quad \sup_{t \in [0, T]} |f(r, t, x_t)| \leq (|f(0, 0, 0)| + K_f(|x_0| + T + d_D) + K_f \lambda^* W_T^*(r)) \exp\{K_f T\},$$

andb

$$\sup_{t \in [0, T]} |f(r, t, x_t)|^2 \leq C_T (|x_0|^2 + 1 + W_T^*(r)^2),$$

with  $W_T^*(r) := \sup_{t \in [0, T]} |W_t(x, r)|$ . Moreover  $W_T^*(r)^2 \leq 2 \left( \sup_{t \in [0, T]} W_t(r) \right)^2 + 2 \left( \sup_{t \in [0, T]} -W_t(r) \right)^2$ , where each of the two terms of the left-hand side of the last sum has the law of  $|W_T(r)|$  under  $P_r$ . Hence, for  $\alpha > 0$

$$\int_{\mathcal{C}} \exp \left\{ \alpha \sup_{t \in [0, T]} |f(r, t, x_t)|^2 \right\} dP_r(x) \stackrel{\text{H\"older}}{\leq} e^{\tilde{\alpha}} \left( \int_{\mathcal{C}} \exp \{3\tilde{\alpha}|x_0|^2\} dP_r(x) \right)^{\frac{1}{3}} \left( \int_{\mathcal{C}} \exp \{6\tilde{\alpha}|W_T(r)|^2\} dP_r(x) \right)^{\frac{2}{3}},$$

where  $\tilde{\alpha} = \alpha C_T$ . For  $\alpha$  small enough,  $\int_{\mathcal{C}} \exp \{3\tilde{\alpha}|x_0|^2\} dP_r(x)$  is finite by hypothesis 7, and  $\int_{\mathcal{C}} \exp \{6\tilde{\alpha}|W_T(r)|^2\} dP_r(x)$  is by (15), so that  $\sup_{t \in [0, T]} |f(r, t, x_t)|^2$  admits exponential moments. Furthermore, the same hypothesis ensures the existence of exponential moments of  $\sup_{t \in [-\tau, 0]} |f(r, t, x_t)|^2$ .

□

## APPENDIX B. PROOF OF LEMMA 2: REGULARITY OF THE SOLUTIONS FOR THE LIMIT EQUATION

In this appendix we demonstrate the regularity in space of the solutions that is expressed in lemma 2. We start by showing a technical lemma on the uncoupled system before proceeding to the proof of that result.

**Lemma 26.** (1) *The map:*

$$\mathcal{P} : \begin{cases} D \rightarrow \mathcal{M}_1^+(\mathcal{C}) \\ r \rightarrow P_r \end{cases}$$

*is continuous with respect to the borel topology on  $D$ , and the weak topology on  $\mathcal{M}_1^+(\mathcal{C})$ , e.g.  $r_n \rightarrow r \implies P_{r_n} \xrightarrow{\mathcal{L}} P_r$ .*

(2) *Let  $\mathcal{W}$  be the Wiener measure on  $\mathcal{C}$ . Then,  $\forall A \in \mathcal{B}(\mathcal{C}), \mathcal{W}(A) = 0 \implies P_r(A) = 0$ .*

(3)  *$P$  is a well defined probability measure on  $\mathcal{C} \times D$ .*

*Proof.* The first point is the consequence of a coupling argument. Let  $W$  be a  $\mathbb{P}$ -Brownian motion,  $\zeta \stackrel{\mathcal{L}}{=} \mu_0$  an initial condition, and  $(r_n)_{n \in \mathbb{N}} \in D^{\mathbb{N}}$  a sequence of positions that converges toward  $r \in D$ . We consider  $X^n$  and  $X$ , the respective strong solutions of the SDEs:

$$\begin{cases} dX_t^n = f(r_n, t, X_t^n)dt + \lambda(r_n)dW_t \\ (X_t^n)_{t \in [-\tau, 0]} = \zeta \end{cases} \quad \begin{cases} dX_t = f(r, t, X_t)dt + \lambda(r)dW_t \\ (X_t)_{t \in [-\tau, 0]} = \zeta \end{cases}$$

driven by the same Brownian motion  $W$ .

Then,

$$\begin{cases} d(X^n - X)_t = (f(r_n, t, X_t^n) - f(r, t, X_t))dt + (\lambda(r_n) - \lambda(r))dW_t \\ ((X^n - X)_t)_{t \in [-\tau, 0]} = 0, \end{cases}$$

so that by Gronwall lemma, letting  $W_T^* = \sup_{t \in [-\tau, T]} |W_t|$ ,

$$\sup_{t \in [-\tau, T]} |(X^n - X)_t| \leq |r_n - r|(K_f T + K_\lambda W_T^*)e^{\{K_f T\}},$$

We conclude by remarking that, for  $\varepsilon > 0$ :

$$\mathbb{P}(\|X^n - X\|_\infty \geq \varepsilon) \leq \mathbb{P}(W_T^* \geq \frac{1}{K_\lambda}(\frac{\varepsilon e^{\{K_f T\}}}{|r_n - r|} - K_f T)),$$

so that  $P_{r_n} = \mathcal{L}(X_n) \implies \mathcal{L}(X) = P_r$  as  $r_n \rightarrow r$ .

In order to prove the second point, let  $W_r$  be the unique strong solution of

$$\begin{cases} dX_t = \lambda(r)dW_t \\ (X_t)_{t \in [-\tau, 0]} = \zeta. \end{cases}$$

Following Exercise (2.10) of [36], we remark, by Lipschitz continuity of  $f$ , that explosion of  $P_r$  almost surely never occurs in finite time, so that Girsanov theorem applies:

$$P_r \ll W_r, \quad \frac{dP_r}{dW_r} = \exp \left\{ \int_0^T \frac{f(r, t, X_t)}{\lambda(r)} dX_t - \frac{1}{2} \int_0^T \left( \frac{f(r, t, X_t)}{\lambda(r)} \right)^2 dt \right\}.$$

Consequently,  $\forall A \in \mathcal{B}(\mathcal{C})$ ,

$$P_r(A) = \mathbb{E}_{W_r} \left( \frac{dP_r}{dW_r} \mathbf{1}_A \right)$$

so that  $P_r(A) = 0$  as soon as  $W_r(A) = 0$ . As  $\lambda(r) > \lambda_*$ ,  $W_r(A) = 0 \iff \mathcal{W}(A) = 0$ .

The third point is now easy to settle. In fact, for any  $y \in \mathcal{C}$  and  $\varepsilon > 0$ ,  $\mathcal{W}(\partial \mathcal{B}(y, \varepsilon)) = \mathcal{W}(\{x \in \mathcal{C}, \|x - y\|_\infty = \varepsilon\}) = 0$ . Hence, Portmanteau implies that  $r \mapsto P_r(\mathcal{B}(y, \varepsilon))$  is a continuous map, so that we can define  $\int_D P_r(\mathcal{B}(y, \varepsilon)) d\pi(r)$  univocally. As  $\{\mathcal{B}(y, \varepsilon) \times B, y \in \mathcal{C}, \varepsilon > 0, B \in \mathcal{B}(D)\}$  form a  $\Pi$ -system that generates  $\mathcal{B}(\mathcal{C} \times D)$ ,  $P$  is a well defined probability measure on  $\mathcal{C} \times D$ .  $\square$

We now proceed to prove lemma 2 that we repeat below:

**Lemma.** *The map*

$$\mathcal{Q} : \begin{cases} D^{\mathbb{N}} \rightarrow \mathcal{M}_1^+(\mathcal{C}^{\mathbb{N}}) \\ \mathbf{r} \rightarrow Q_{\mathbf{r}}^{\mathbb{N}} \end{cases}$$

where  $Q_{\mathbf{r}}^N := \mathcal{E}_J(Q_{\mathbf{r}}^N(J))$ , is continuous with respect to the weak topology. Moreover,

$$dQ^N(\mathbf{x}, \mathbf{r}) := dQ_{\mathbf{r}}^N(\mathbf{x}) d\pi^{\otimes N}(\mathbf{r})$$

defines a probability measure on  $\mathcal{M}_1^+(\mathcal{C} \times D)^N$ .

*Remark 9.*  $\mathcal{Q}$  maps the positions  $\mathbf{r}$  to the Gaussian averaged of the solutions  $Q_{\mathbf{r}}^N(J)$ , so that its continuity seems to be a consequence of Cauchy-Lipschitz theorem with parameter  $\mathbf{r}$ . Yet, the equation depends on  $\mathbf{r}$  through the synaptic weights  $J_{ij}$  which only satisfy a continuity in law. Meanwhile the proof is not difficult, it must rely on another argument. The one developped here is a coupling method.

*Proof.* Let  $\mathbf{r}^n \rightarrow \mathbf{r} \in D^N$ , let  $(W_t^i, 0 \leq t \leq T)_{i \in \llbracket 1, N \rrbracket}$  be a family of independent  $\mathbb{P}$ -Brownian motions, and  $\zeta \stackrel{\mathcal{L}}{=} \mu_0$  an initial condition. Let now  $X_{\mathbf{r}^n}^N = (X_{\mathbf{r}^n}^{i,N})_{i \in \llbracket 1, N \rrbracket}$  and  $X_{\mathbf{r}}^N = (X_{\mathbf{r}}^{i,N})_{i \in \llbracket 1, N \rrbracket}$  be the respective strong solutions of the two following stochastic differential equations:

$$\begin{cases} dX_{\mathbf{r}^n}^{i,N}(t) = \left( f(r_i^n, t, X_{\mathbf{r}^n}^{i,N}(t)) + \sum_{j=1}^N \tilde{J}_{ij}^{\mathbf{r}^n} S(X_{\mathbf{r}^n}^{j,N}(t - \tau_{r_i^n r_j^n})) \right) dt + \lambda(r_i^n) dW_t^i \\ (X_{\mathbf{r}^n}^N(t))_{t \in [-\tau, 0]} = \zeta^{\otimes N}, \end{cases}$$

$$\begin{cases} dX_{\mathbf{r}}^{i,N}(t) = \left( f(r_i, t, X_{\mathbf{r}}^{i,N}(t)) + \sum_{j=1}^N J_{ij}^{\mathbf{r}} S(X_{\mathbf{r}}^{j,N}(t - \tau_{r_i r_j})) \right) dt + \lambda(r_i) dW_t^i \\ (X_{\mathbf{r}}^N(t))_{t \in [-\tau, 0]} = \zeta^{\otimes N}. \end{cases}$$

where  $J_{ij}^{\mathbf{r}} \sim \mathcal{N}\left(\frac{J(r_i, r_j)}{N}, \frac{\sigma(r_i, r_j)^2}{N}\right)$ ,  $\tilde{J}_{ij}^{\mathbf{r}^n} \sim J_{ij}^{\mathbf{r}^n}$  satisfy (5), and where we used the short-hand notation  $\tau_{rr'} := \tau(r, r')$ . In particular,  $X_{\mathbf{r}^n}^{i,N}$  has law  $Q_{\mathbf{r}^n}^N(J^{\mathbf{r}^n})$ , and  $X_{\mathbf{r}}^{i,N}$  has law  $Q_{\mathbf{r}}^N(J^{\mathbf{r}})$ .

Then,

$$\begin{cases} d(X_{\mathbf{r}^n}^{i,N} - X_{\mathbf{r}}^{i,N})(t) = \left( f(r_i^n, t, X_{\mathbf{r}^n}^{i,N}(t)) - f(r_i, t, X_{\mathbf{r}}^{i,N}(t)) \right) dt \\ + \sum_{j=1}^N \left( \tilde{J}_{ij}^{\mathbf{r}^n} S(X_{\mathbf{r}^n}^{j,N}(t - \tau_{r_i^n r_j^n})) - J_{ij}^{\mathbf{r}} S(X_{\mathbf{r}}^{j,N}(t - \tau_{r_i r_j})) \right) dt + (\lambda(r_i^n) - \lambda(r_i)) dW_t^i, \\ ((X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N)_t)_{t \in [-\tau, 0]} = 0, \end{cases}$$

so that

$$\begin{aligned} (X_{\mathbf{r}^n}^{i,N} - X_{\mathbf{r}}^{i,N})(t) &= \left( \int_0^t \left( f(r_i^n, s, X_{\mathbf{r}^n}^{i,N}(s)) - f(r_i, s, X_{\mathbf{r}}^{i,N}(s)) \right) ds \right. \\ &\quad \left. + \sum_{j=1}^N \left\{ (\tilde{J}_{ij}^{\mathbf{r}^n} - J_{ij}^{\mathbf{r}}) \int_0^t S(X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n})) ds \right. \right. \\ &\quad \left. \left. + J_{ij}^{\mathbf{r}} \int_0^t \left( S(X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n})) - S(X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j})) \right) ds \right\} \right) + (\lambda(r_i^n) - \lambda(r_i)) W_t^i. \end{aligned}$$

Let  $W_T^* = \sup_{t \in [-\tau, T], 1 \leq i \leq N} |W_t^i|$ . This quantity is almost surely finite under  $\mathbb{P}$ . Using Lipschitz continuity of  $f$ ,  $\lambda$ , and the fact that  $\sup |S| \leq 1$ , one obtains

$$\begin{aligned} |X_{\mathbf{r}^n}^{i,N} - X_{\mathbf{r}}^{i,N}|(t) &\leq \left( \int_0^t K_f (|r_i^n - r_i| + |X_{\mathbf{r}^n}^{i,N}(s) - X_{\mathbf{r}}^{i,N}(s)|) ds + \sum_{j=1}^N \left\{ t |\tilde{J}_{ij}^{\mathbf{r}^n} - J_{ij}^{\mathbf{r}}| \right. \right. \\ &\quad \left. \left. + |J_{ij}^{\mathbf{r}}| \int_0^t |S(X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n})) - S(X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j}))| ds \right\} \right) + K_\lambda |r_i^n - r_i| W_T^*. \end{aligned}$$

Moreover,

$$\begin{aligned} &|J_{ij}^{\mathbf{r}}| |S(X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n})) - S(X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j}))| \\ &\leq M |S(X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n})) - S(X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j}))| + 2 |J_{ij}^{\mathbf{r}}| \mathbf{1}_{|J_{ij}^{\mathbf{r}}| > M} \\ &\leq M K_S \left( |X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n}) - X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j})| + |X_{\mathbf{r}^n}^{j,N} - X_{\mathbf{r}}^{j,N}|(s - \tau_{r_i^n r_j^n}) \right) + 2 |J_{ij}^{\mathbf{r}}| \mathbf{1}_{|J_{ij}^{\mathbf{r}}| > M}. \end{aligned}$$

taking the supremum in  $i \in \{1, \dots, N\}$  yields

$$\begin{aligned} &\|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(t) \leq \|\mathbf{r}^n - \mathbf{r}\|_\infty (TK_f + K_\lambda W_T^*) \\ &+ \sum_{i,j=1}^N \left\{ T |\tilde{J}_{ij}^{\mathbf{r}^n} - J_{ij}^{\mathbf{r}}| + M K_S \int_0^t |X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n}) - X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j})| ds + 2 |J_{ij}^{\mathbf{r}}| \mathbf{1}_{|J_{ij}^{\mathbf{r}}| > M} \right\} \\ &+ K_f \int_0^t \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(s) ds + N M K_S \int_0^t \sup_{-\tau \leq u \leq s} \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(u) ds. \end{aligned}$$

We can now take the supremum in time

$$\begin{aligned} &\sup_{s \in [-\tau, t]} \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(s) \leq C_T \left\{ \|\mathbf{r}^n - \mathbf{r}\|_\infty (1 + W_T^*) + (1 + MN) \int_0^t \sup_{u \in [-\tau, s]} \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(u) ds \right. \\ &+ \sum_{i,j=1}^N \left\{ |\tilde{J}_{ij}^{\mathbf{r}^n} - J_{ij}^{\mathbf{r}}| + |J_{ij}^{\mathbf{r}}| \mathbf{1}_{|J_{ij}^{\mathbf{r}}| > M} + M \int_0^t |X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n}) - X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j})| ds \right\} \Bigg\}, \end{aligned}$$

and integrate on  $J$  to obtain by condition (5) and hypothesis  $(H_J)$

$$\begin{aligned} \mathcal{E}_J \left[ \sup_{s \in [-\tau, t]} \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(s) \right] &\leq C_T \left\{ \|\mathbf{r}^n - \mathbf{r}\|_\infty (N + W_T^*) + N^2 M D_0 \exp\{-aNM^2\} \right. \\ &\quad \left. + MN^2 m_n + (1 + MN) \int_0^t \mathcal{E}_J \left[ \sup_{u \in [-\tau, s]} \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(u) \right] ds \right\}. \end{aligned}$$

where  $m_n := \max_{i,j \in \llbracket 1, N \rrbracket} \int_0^T \mathcal{E}_J \left[ |X_{\mathbf{r}^n}^{j,N}(s - \tau_{r_i^n r_j^n}) - X_{\mathbf{r}}^{j,N}(s - \tau_{r_i r_j})| \right] ds$ . As solution are  $\mathbb{P}$  almost surely continuous, and  $N$  remains (here) finite, this quantity tends to zero with probability one when  $n$  goes to infinity. Moreover, Gronwall's lemma ensures

(44)

$$\mathcal{E}_J \left[ \sup_{t \in [-\tau, T]} \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(t) \right] \leq N^2 C_T \left( \|\mathbf{r}^n - \mathbf{r}\|_\infty (1 + W_T^*) + M(e^{-aNM^2} + m_n) \right) e^{C_T(1+MN)}.$$

Choosing  $M_n = \sqrt{-\log(\max(\|\mathbf{r}^n - \mathbf{r}\|_\infty, m_n))}$ , one sees that  $\mathcal{E}_J \left[ \sup_{t \in [-\tau, T]} \|X_{\mathbf{r}^n}^N - X_{\mathbf{r}}^N\|_\infty^{\mathbb{R}^N}(t) \right] \rightarrow 0$   $\mathbb{P}$ -a.s. as  $n \rightarrow +\infty$ . Hence, for any  $\phi : \mathcal{C}^N \rightarrow \mathbb{R}$  Lipschitz continuous bounded function, the Dominated Convergence Theorem ensures that

$$\mathcal{E}_J \left[ \int_{\Omega} \phi(X_{\mathbf{r}^n}^N(\omega)) d\mathbb{P}(\omega) \right] \xrightarrow{n \rightarrow \infty} \mathcal{E}_J \left[ \int_{\Omega} \phi(X_{\mathbf{r}}^N(\omega)) d\mathbb{P}(\omega) \right].$$

In other words,  $Q_{\mathbf{r}^n}^N \implies Q_{\mathbf{r}}^N$ , and the map  $\mathbf{r} \rightarrow \mathcal{E}_J \left[ \int_{\mathcal{C}^N} \phi(\mathbf{x}) dQ_{\mathbf{r}}^N(\mathbf{x}) \right]$  is continuous and integrable with respect to  $\pi^{\otimes N}$ . In particular,  $dQ^N(\mathbf{x}, \mathbf{r}) := dQ_{\mathbf{r}}^N(\mathbf{x}) d\pi^{\otimes N}(\mathbf{r})$  defines a probability measure on  $(\mathcal{C} \times D)^N$ .  $\square$

#### APPENDIX C. PROOF OF LEMMA 14: APPROXIMATION OF $\Gamma$ BY $\Gamma_\nu$

Proof of Lemma 14: (1):

*Proof.* Using Hölder inequality with conjugate exponents  $(\sigma, \eta)$ , one finds:

$$(45) \quad B^N \leq \left\{ \mathcal{E} \left[ \int \exp \left\{ \sum_{i=1}^N \sigma X_i^\nu \right\} dP^{\otimes N} \right] \right\}^{\frac{1}{\sigma}} \underbrace{\left\{ \mathcal{E} \left[ \int \mathbb{1}_{\{\hat{\mu}_N \in B(\nu, \delta)\} \cap A_{N, \delta}^1} \prod_{i=1}^N \exp a\eta (X_i^{\hat{\mu}_N} - X_i^\nu) dP^{\otimes N} \right] \right\}^{\frac{1}{\eta}}}_{B_2^N}.$$

The first term is controlled by martingale property:

$$\begin{aligned} B_1^N &= \mathcal{E} \left[ \int_{D^N} d\pi^{\otimes N}(\mathbf{r}) \exp \left\{ \sum_{i=1}^N \frac{\sigma^2 - \sigma}{2} \int_0^T (G_t^\nu(r_i) + m_\nu(t, r_i))^2 dt \right\} \right. \\ &\quad \times \left. \int_{\mathcal{C}^N} \exp \left\{ \sum_{i=1}^N \sigma \int_0^T (G_t^\nu(r_i) + m_\nu(t, r_i)) dW_t(x^i, r_i) - \frac{\sigma^2}{2} \int_0^T (G_t^\nu(r_i) + m_\nu(t, r_i))^2 dt \right\} dP_{\mathbf{r}}(\mathbf{x}) \right] \\ &\stackrel{\text{Jensen}}{\leq} \int_{D^N} \prod_{i=1}^N \int_0^T \mathcal{E} \left[ \exp \left\{ \frac{\sigma(\sigma-1)T}{2} (G_t^\nu(r_i) + m_\nu(t, r_i))^2 \right\} \right] \frac{dt}{T} d\pi^{\otimes N}(\mathbf{r}) \stackrel{(15)}{\leq} \exp \{c_T(\sigma-1)N\}, \end{aligned}$$

with  $c_T$  uniform in space.

The second term necessitates a fine control of space regularity of the solution that we now perform. We suppose that  $\delta$  is arbitrarily small, and set  $\kappa = a\eta$ . By

Cauchy-Schwarz inequality:

$$\begin{aligned}
B_2^N &\leq \left\{ \mathcal{E} \left[ \int \prod_{i=1}^N \exp \left\{ 2\kappa \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i) + (m_{\hat{\mu}_N} - m_\nu)(t, r_i)) dW_t(x^i, r_i) \right. \right. \right. \\
&\quad \left. \left. \left. - 2\kappa^2 \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i) + (m_{\hat{\mu}_N} - m_\nu)(t, r_i))^2 dt \right\} dP^{\otimes N} \right] \right\}^{\frac{1}{2}} \\
&\times \left\{ \int_{\hat{\mu}_N \in B(\nu, \delta)} \mathcal{E} \left[ \mathbb{1}_{A_{N, \delta}^1} \prod_{i=1}^N \exp \left\{ 2\kappa^2 \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i) + (m_{\hat{\mu}_N} - m_\nu)(t, r_i))^2 dt \right. \right. \right. \\
&\quad \left. \left. \left. - \kappa \int_0^T (G_t^{\hat{\mu}_N}(r_i) + m_{\hat{\mu}_N}(t, r_i))^2 - (G_t^\nu(r_i) + m_\nu(t, r_i))^2 dt \right\} dP^{\otimes N} \right] \right\}^{\frac{1}{2}}
\end{aligned}$$

The first term is bounded by one by supermartingale properties. For the second term, we remark that:

$$\begin{aligned}
& - \int_0^T (G_t^{\hat{\mu}_N}(r_i) + m_{\hat{\mu}_N}(t, r_i))^2 - (G_t^\nu(r_i) + m_\nu(t, r_i))^2 dt \leq \\
& \quad \frac{\delta^{\frac{1}{2}}}{2} \left( \frac{1}{\delta} \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i) + (m_{\hat{\mu}_N} - m_\nu)(t, r_i))^2 dt \right. \\
& \quad \left. + \int_0^T (G_t^{\hat{\mu}_N}(r_i) + G_t^\nu(r_i) + (m_{\hat{\mu}_N} + m_\nu)(t, r_i))^2 dt \right)
\end{aligned}$$

so that, by Cauchy-Schwarz inequality:

$$\begin{aligned}
B_2^N &\leq \left\{ \int_{\hat{\mu}_N \in B(\nu, \delta)} \mathcal{E} \left[ \mathbb{1}_{A_{N, \delta}^1} \prod_{i=1}^N \exp \left\{ (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i) + (m_{\hat{\mu}_N} - m_\nu)(t, r_i))^2 dt \right\} dP^{\otimes N} \right] \right\}^{\frac{1}{4}} \\
&\times \left\{ \int \prod_{i=1}^N \underbrace{\mathcal{E} \left[ \exp \left\{ \kappa\delta^{\frac{1}{2}} \int_0^T (G_t^{\hat{\mu}_N}(r_i) + G_t^\nu(r_i) + (m_{\hat{\mu}_N} + m_\nu)(t, r_i))^2 dt \right\} dP^{\otimes N} \right]}_{\stackrel{(15)}{\leq} \exp \left\{ c_T \kappa \delta^{\frac{1}{2}} \right\}} \right\}^{\frac{1}{4}}.
\end{aligned}$$

The first term of the product is more intricate. For  $\hat{\mu}_N \in B(\nu, \delta)$ :

$$\begin{aligned}
|m_{\hat{\mu}_N} - m_\nu|(t, r_i) &= \frac{1}{\lambda_*} \int |J(r_i, r') S(x_{t-\tau(r_i, r')}) - J(r_i, \tilde{r}') S(y_{t-\tau(r_i, \tilde{r}')})| d\xi((x, r'), (y, \tilde{r}')) \\
&\leq \frac{KJ}{\lambda_*} \int |r' - \tilde{r}'| d\xi + \frac{\|\bar{J}\|_\infty}{\lambda_*} \int (\mathbb{1}_{\{|r' - \tilde{r}'| \leq \delta\}} + \mathbb{1}_{\{|r' - \tilde{r}'| > \delta\}}) |S(x_{t-\tau(r_i, r')}) - S(y_{t-\tau(r_i, \tilde{r}')})| d\xi \\
&\leq C \left\{ (1 + \delta^{-1}) \inf_\xi \int |r' - \tilde{r}'| d\xi + \inf_\xi \int \mathbb{1}_{\{|r' - \tilde{r}'| \leq \delta\}} |S(x_{t-\tau(r_i, r')}) - S(y_{t-\tau(r_i, \tilde{r}')})| d\xi \right\} \\
&\leq C\delta + \frac{C}{N} \sum_{j=1}^N \int \mathbb{1}_{\{|r_j - r'| \leq \delta\}} |S(x_{t-\tau(r_i, r_j)}^j) - S(x_{t-\tau(r_i, r')}^j)| d\nu(y, r'),
\end{aligned}$$

where we have, in the last inequality, dominated the infimum with the particular measure  $\xi = \hat{\mu}_N \otimes \nu$ . We will now take advantage of being on  $\{\hat{\mu}_N \in B(\nu, \delta)\} \cap A_{N,\delta}^1$ :

$$\begin{aligned} |m_{\hat{\mu}_N} - m_\nu|(t, r_i) &\leq C\delta + \frac{C}{N} \int \left\{ \sum_{j=1}^N \mathbb{1}_{\{|r_j - r'| \leq \delta\}} \mathbb{1}_{E_N^j} K_S |x_{t-\tau(r_i, r_j)}^j - x_{t-\tau(r_i, r')}^j| \right\} d\nu(y, r') + \frac{2C}{N} c_N \\ &\leq C\delta^{2\beta} + \frac{C}{N} \int \left\{ \sum_{j=1}^N \mathbb{1}_{\{|r_j - r'| \leq \delta\}} \mathbb{1}_{E_N^j} \delta^{2\beta} \right\} d\nu(y, r') \leq C\delta^{2\beta}. \end{aligned}$$

Hence, Jensen inequality gives

$$\left( G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i) + (m_{\hat{\mu}_N} - m_\nu)(t, r_i) \right)^2 \leq C_T \left[ \delta^{2\beta} + (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i))^2 \right].$$

Then

$$\begin{aligned} &\mathcal{E} \left[ \mathbb{1}_{A_{N,\delta}^1} \prod_{i=1}^N \exp \left\{ (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i) + (m_{\hat{\mu}_N} - m_\nu)(t, r_i))^2 dt \right\} \right] \leq \\ (46) \quad &\exp \left\{ C_T (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) \delta^{2\beta} N \right\} \mathcal{E} \left[ \mathbb{1}_{A_{N,\delta}^1} \prod_{i=1}^N \exp \left\{ C_T (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i))^2 dt \right\} \right]. \end{aligned}$$

To bound the last term of (46), we proceed as precedently to find on  $A_{N,\delta}^1$ ,

$$\begin{aligned} \mathcal{E} \left[ (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i))^2 \right] &= \int_{(\mathcal{C} \times D)^2} \frac{1}{\lambda(r_i)^2} |\sigma(r_i, r') S(x_{t-\tau(r_i, r')}) - \sigma(r_i, \tilde{r}') S(y_{t-\tau(r_i, \tilde{r}')})|^2 d\xi \\ &\leq C_T \left\{ (1 + d_D \delta^{-1}) d_T (\hat{\mu}_N, \nu)^2 + \frac{1}{N} \sum_{j=1}^N \int \mathbb{1}_{\{|r_j - r'| \leq \delta\}} |S(x_{t-\tau(r_i, r_j)}^j) - S(x_{t-\tau(r_i, r')}^j)|^2 d\nu(y, r') \right\} \\ &\leq C_T \left\{ \delta + \frac{1}{N} \int \left\{ \sum_{j=1}^N \mathbb{1}_{\{|r_j - r'| \leq \delta\}} \mathbb{1}_{E_N^j} \delta^{2\beta} \right\} d\nu(y, r') + \frac{C}{N} c_N \right\} \leq C\delta^{2\beta}. \end{aligned}$$

Now

$$\begin{aligned} &\mathcal{E} \left[ \mathbb{1}_{A_{N,\delta}^1} \prod_{i=1}^N \exp \left\{ C_T (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) \int_0^T (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i))^2 dt \right\} \right] \\ &\stackrel{\text{Jensen}}{\leq} \int_0^T \mathcal{E} \left[ \mathbb{1}_{A_{N,\delta}^1} \prod_{i=1}^N \exp \left\{ C_T (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) (G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i))^2 \right\} \right] \frac{dt}{T} \\ &\leq \int_0^T \prod_{i=1}^N \mathcal{E} \left[ \exp \left\{ C_T (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) \delta^{2\beta} \frac{(G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i))^2}{\mathcal{E}[(G_t^{\hat{\mu}_N}(r_i) - G_t^\nu(r_i))^2]} \right\} \right] \frac{dt}{T} \stackrel{(15)}{\leq} \exp \{ C_T (4\kappa^2 + \kappa\delta^{-\frac{1}{2}}) \delta^{2\beta} N \}, \end{aligned}$$

by independence of the  $G$  for different locations. Hence,

$$B_N^2 \leq \exp \{ C_\kappa(\delta) N \}$$

with  $C_\kappa(\delta) \rightarrow 0$  as  $\delta \rightarrow 0$ .

(2): Under a slightly different short-time hypothesis due to the presence of Gaussian synaptic weights  $\frac{2T\|\sigma\|_\infty^2}{\lambda_*^2} < 1$ , the proof proceed exactly as in (23).  $\square$



## APPENDIX D. NON-GAUSSIAN ESTIMATES

We prove that the different assumptions (H1)-(H4) are valid.

(H1): By a direct application of Jensen inequality

$$\begin{aligned} a_i^\delta(\mathbf{x}, \mathbf{r}) &\geq \exp \left\{ \int_0^t m_{\hat{\mu}_N}(t^{(l)}, r_i) dW_t(x^i, r_i) - \frac{1}{2} \int_0^T (m_{\hat{\mu}_N}(t^{(l)}, r_i)^2 + K_{\hat{\mu}_N}(t^{(l)}, t^{(l)}, r_i)) dt \right\} \\ &\geq \exp \left\{ -\frac{1}{2} \left( \frac{\|\bar{J}\|_\infty^2 + \|\sigma\|_\infty^2}{\lambda_*^2} \right) T \right\} \exp \left\{ -\frac{\|\bar{J}\|_\infty \sqrt{\delta T}}{\lambda_*} \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_i, r_i)| \right\}. \end{aligned}$$

(H2): Remark that

$$\begin{aligned} a_i^\delta(\mathbf{x}, \mathbf{r}) &= \exp \left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}(x_i, r_i)^2 \right\} \mathcal{E}_J \left( \exp \left\{ -\frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left( \sqrt{\delta T} \hat{G}_{t^{(l)}}^{\delta, i}(\mathbf{x}, \mathbf{r}) - B_{t_l}(x^i, r_i) \right)^2 \right\} \right), \\ &\stackrel{\text{H\"older}}{\leq} \exp \left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}(x_i, r_i)^2 \right\} \prod_{l=1}^{\frac{1}{\delta}} \mathcal{E}_J \left( \exp \left\{ -\frac{1}{2\delta} \left( \sqrt{\delta T} \hat{G}_{t^{(l)}}^{\delta, i}(\mathbf{x}, \mathbf{r}) - B_{t_l}(x^i, r_i) \right)^2 \right\} \right)^\delta. \end{aligned}$$

Suppose first that  $B_{t_l}(x^i, r_i) \geq 0$ . Then

$$\begin{aligned} \left( B_{t_l}(x_i, r_i) - \sqrt{\delta T} \hat{G}_{t^{(l)}}^{\delta, i}(\mathbf{x}, \mathbf{r}) \right)^2 &\geq \left( B_{t_l}(x_i, r_i) - \sqrt{\delta T} \hat{G}_{t^{(l)}}^{\delta, i}(\mathbf{x}, \mathbf{r}) \right)^2 \mathbf{1}_{\{2\sqrt{\delta T} \hat{G}_{t^{(l)}}^{\delta, i} \leq B_{t_l}\}} \\ &\geq \frac{B_{t_l}(x_i, r_i)^2}{4} \mathbf{1}_{\{2\sqrt{\delta T} \hat{G}_{t^{(l)}}^{\delta, i} \leq B_{t_l}\}}, \end{aligned}$$

so that

$$\mathcal{E}_J \left( \exp \left\{ -\frac{1}{2\delta} \left( \sqrt{\delta T} \hat{G}_{t^{(l)}}^{\delta, i}(\mathbf{x}, \mathbf{r}) - B_{t_l} \right)^2 \right\} \right) \stackrel{(H_J)}{\leq} \exp \left\{ -\frac{B_{t_l}^2}{8\delta} \right\} + D_0 \exp \left\{ -\frac{\lambda_*^2 a}{4\delta T} B_{t_l}^2 \right\}.$$

We obtain the same inequality under the hypothesis  $B_{t_l} \leq 0$ , so that, making use of hypothesis (H<sub>J</sub>)

$$a_i^\delta(\mathbf{x}, \mathbf{r}) \leq \max(1, D_0) \exp \left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}(x_i, r_i)^2 \right\} \prod_{l=1}^{\frac{1}{\delta}} \exp \left\{ -\frac{1}{4} \min \left( \frac{1}{2}, \frac{\lambda_*^2 a}{T} \right) B_{t_l}^2 \right\}.$$

(H3): Following the exact proof of Moynot and Samuelides [32], with constants  $q = \frac{a\lambda_*^2}{\sqrt{2T}} \sqrt{\frac{N}{k}}$ , and  $\epsilon = \sqrt{2\frac{k}{N}}$ , we obtain

$$\tilde{a}_i^\delta(\mathbf{x}, \mathbf{r}) \leq \exp \left\{ \left( \frac{1}{2q} + \frac{a\lambda_*^2}{2Tq} \right) \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right\} \left( \frac{D_0}{A} \right)^{\frac{1}{q}} \exp \left\{ \frac{B\sqrt{\delta}}{q} \sum_{l=1}^{\frac{1}{\delta}} |B_{t_l}(x_i, r_i)| \right\}.$$

Remark that  $\sum_{l=1}^{\frac{1}{\delta}} \sqrt{\delta} |B_{t_l}(x_i, r_i)| \leq \frac{1}{2} (1 + \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i))$ , so that taking  $q$  large enough, i.e.  $\frac{k}{N}$  small enough, yields the result.

(H4):

$$b_i^\delta(\mathbf{x}, \mathbf{r}) = \exp \left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right\} \mathcal{E}_J \left( \exp \left\{ -\frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left( \sqrt{\delta T} \tilde{G}_{t^{(l)}}^{\delta, i}(\mathbf{x}, \mathbf{r}) - B_{t_l}(x_i, r_i) \right)^2 \right\} \right).$$

Under  $\mathcal{E}_J$ ,  $U_i^\delta := \left( \sqrt{\delta T} \tilde{G}_{t^{(l)}}^{\delta, i} \right)_{l \in \llbracket 1, \frac{1}{\delta} \rrbracket}$  is a Gaussian vector of mean  $\bar{U}_i^\delta := \left( \sqrt{\delta T} m_{\hat{\mu}_N}(t^{(l)}, r_i) \right)_{l \in \llbracket 1, \frac{1}{\delta} \rrbracket}$  and variance-covariance matrix  $\Sigma_i^\delta := \left( \delta T K_{\hat{\mu}_N}(t_l, t_m, r_i) \right)_{(l, m) \in \llbracket 1, \frac{1}{\delta} \rrbracket^2}$ . Let

$$(47) \quad A_{N, \delta}^1 := \{(\mathbf{x}, \mathbf{r}) \in (\mathcal{C} \times D)^N, \forall Y \in \mathbb{R}^{\frac{1}{\delta}} \setminus \{0\}, \forall i \in \llbracket 1, N \rrbracket, \exists j \in \llbracket 1, N \rrbracket, \left( \sum_{l=1}^{\frac{1}{\delta}} y_l S(x_{t_l - \tau(r_i, r_j)}^j) \right)^2 > 0\}.$$

As

$${}^t Y \Sigma_i^\delta Y = \frac{\delta T}{\lambda(r_i)^2 N} \sum_{j=1}^N \sigma(r_i, r_j)^2 \left( \sum_{l=1}^{\frac{1}{\delta}} y_l S(x_{t_l - \tau(r_i, r_j)}^j) \right)^2,$$

the matrix  $\Sigma_i^\delta$  is positive definite on  $A_{N, \delta}^2$ . Being on this set basically mean that the  $N$  random vectors  $(S(x_{t_l - \tau(r_i, r_j)}^j))_{l \in \llbracket 1, \frac{1}{\delta} \rrbracket}$ ,  $1 \leq j \leq N$  are not contained in an hyperplane of  $\mathbb{R}^{\frac{1}{\delta}}$ . For  $N \geq \frac{1}{\delta}$ ,  $P^{\otimes N}(A_{N, \delta}^2) = 1$  as, under  $P_{\mathbf{r}}$ , the  $x^j$  are independent semi-martingales. Let  $B := (B_{t_l})_{l \in \llbracket 1, \frac{1}{\delta} \rrbracket}$ . Hence,

$$\begin{aligned} b_i^\delta(\mathbf{x}, \mathbf{r}) &= \exp \left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right\} \mathcal{E}_J \left( \exp \left\{ -\frac{1}{2} (U_i^\delta - \bar{U}_i^\delta + \bar{U}_i^\delta - B) \cdot (U_i^\delta - \bar{U}_i^\delta + \bar{U}_i^\delta - B) \right\} \right), \\ b_i^\delta(\mathbf{x}, \mathbf{r}) &= e^{\left\{ \frac{1}{2} B \cdot B \right\}} e^{\left\{ -\frac{1}{2} (\bar{U}_i^\delta - B) \cdot (\bar{U}_i^\delta - B) \right\}} \int e^{\left\{ -X \cdot (\bar{U}_i^\delta - B) \right\}} \frac{e^{\left\{ -\frac{1}{2} X \cdot (I_{\frac{1}{\delta}} + (\Sigma_i^\delta)^{-1}) \cdot X \right\}}}{\sqrt{(2\pi)^N \det(\Sigma_i^\delta)}} dX, \\ &= \sqrt{\det \left( I_{\frac{1}{\delta}} + \Sigma_i^\delta \right)^{-1}} \exp \left\{ \frac{1}{2} B \cdot B \right\} \exp \left\{ \frac{1}{2} (\bar{U}_i^\delta - B) \cdot (A_i^\delta - I_{\frac{1}{\delta}}) \cdot (\bar{U}_i^\delta - B) \right\}, \end{aligned}$$

where  $A_i^\delta = \Sigma_i^\delta \left( I_{\frac{1}{\delta}} + \Sigma_i^\delta \right)^{-1}$  is a definite postive matrix with eigenvalues strictly smaller than 1. Consequently,

$$(b_i^\delta)^{-1} \leq \underbrace{\sqrt{\det \left( I_{\frac{1}{\delta}} + \Sigma_i^\delta \right)} \exp \left\{ \frac{1}{2} (\bar{U}_i^\delta - B) \cdot (\bar{U}_i^\delta - B) \right\} \exp \left\{ -\frac{1}{2} B \cdot B \right\}}_{N_i}.$$

As

$$|a_i^\delta - b_i^\delta| = e^{\left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} B_{t_l}^2(x_i, r_i) \right\}} \underbrace{\left| \mathcal{E}_J \left( e^{\left\{ -\frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left( \sqrt{\delta T} \tilde{G}_{t^{(l)}}^{\delta, i} - B_{t_l} \right)^2 \right\}} - e^{\left\{ -\frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left( \sqrt{\delta T} \tilde{G}_{t^{(l)}}^{\delta, i} - B_{t_l} \right)^2 \right\}} \right) \right|}_{O_i},$$

then

$$\left| \frac{a_i^\delta - b_i^\delta}{b_i^\delta} \right| \leq O_i N_i.$$

Remark that

$$\sqrt{\det \left( I_{\frac{1}{\delta}} + \Sigma_i^\delta \right)} = \exp \left\{ \frac{1}{2} \sum_{\lambda \in sp(\Sigma_i^\delta)} \log(1 + \lambda) \right\} \leq \exp \left\{ \frac{1}{2} tr(\Sigma_i^\delta) \right\} \leq \exp \left\{ \frac{\|\sigma\|_\infty^2 T}{2\lambda_*^2} \right\},$$

so that

$$N_i \leq \exp \left\{ T \frac{\|\sigma\|_\infty^2 + \|\bar{J}\|_\infty^2}{2\lambda_*^2} \right\} \exp \left\{ \frac{1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left( B_{t_l}^2 + 2\sqrt{\delta T} \frac{\|\bar{J}\|_\infty}{\lambda_*} |B_{t_l}| \right) \right\}.$$

To obtain an upperbound for  $O_i$  we rely on [32, Lemma 4.2] where the fixed and finite  $\frac{1}{\delta}$  correspond to their  $T$ . Following its proof, we define the function:

$$\Phi(y_1, \dots, y_{\frac{1}{\delta}}) = \prod_{l=1}^{\frac{1}{\delta}} \phi(y_l + a_l),$$

where  $\phi(z) = \exp -\frac{z^2}{2}$  and  $a_l = \frac{\sqrt{\delta T}}{\lambda(r_i)} \sum_{j=1}^N \frac{J(r_i, r_j)}{N} S(x_{t^{(l)} - \tau(r_i, r_j)}^j) - B_{t_l}$ . One easily sees that the three first derivative of  $\Phi$  are bounded by some constant  $C_3$  independent of  $\delta$ . Let

$$(Y_j)_l = \frac{\sqrt{\delta T}}{\lambda(r_i)} \left( J_{ij} - \frac{J(r_i, r_j)}{N} \right) S(x_{t^{(l)} - \tau(r_i, r_j)}^j)$$

$$(\tilde{Y}_j)_l = \frac{\sqrt{\delta T}}{\lambda(r_i)} \left( \tilde{J}_{ij} - \frac{J(r_i, r_j)}{N} \right) S(x_{t^{(l)} - \tau(r_i, r_j)}^j),$$

so that

$$O_i = \left| \mathcal{E}_J \left( \Phi \left( \sum_{j=1}^N Y_j \right) \right) - \Phi \left( \sum_{j=1}^N \tilde{Y}_j \right) \right|.$$

Let  $\varepsilon > 0$ . Then

$$O_i \leq C_3 \left( \frac{\varepsilon}{6} \sum_{j=1}^N \mathcal{E}_J \left( \|Y_j\|^2 \right) + \sum_{j=1}^N \mathcal{E}_J \left( \|\tilde{Y}_j\|^3 \right) + \sum_{j=1}^N \mathcal{E}_J \left( \|Y_j\|^2 \mathbf{1}_{\{\|Y_j\| \geq \varepsilon\}} \right) \right),$$

where  $\|X\| = \sqrt{\sum_{l=1}^{\frac{1}{\delta}} x_l^2}$ . But

$$\mathcal{E}_J \left( \|Y_j\|^2 \right) \leq \frac{\delta T}{\lambda_*^2} \mathcal{E}_J \left[ \left( J_{ij} - \frac{J(r_i, r_j)}{N} \right)^2 \right] \sum_{l=1}^{\frac{1}{\delta}} S(x_{t^{(l)} - \tau(r_i, r_j)}^j)^2 \leq \frac{\|\sigma\|_\infty^2 T}{N \lambda_*^2},$$

$$\mathcal{E}_J \left( \|\tilde{Y}_j\|^3 \right) = \left( \frac{\sqrt{\delta T}}{\lambda(r_i)} \sqrt{\sum_{l=1}^{\frac{1}{\delta}} S(x_{t^{(l)} - \tau(r_i, r_j)}^j)^2} \right)^3 \mathcal{E}_J \left( \left| \tilde{J}_{ij} - \frac{J(r_i, r_j)}{N} \right|^3 \right) \leq \frac{\|\sigma\|_\infty^3 T^{\frac{3}{2}}}{N^{\frac{3}{2}} \lambda_*^3} \mathcal{E}_J \left( |\mathcal{N}(0, 1)|^3 \right),$$

$$\mathcal{E}_J \left( \|Y_j\|^2 \mathbf{1}_{\{\|Y_j\| \geq \varepsilon\}} \right) \leq \frac{1}{N} \mathcal{E}_J \left( \frac{\delta T N}{\lambda_*^2} \left( J_{ij} - \frac{J(r_i, r_j)}{N} \right)^2 \mathbf{1}_{\left\{ \frac{\delta T}{\lambda_*^2} \left( J_{ij} - \frac{J(r_i, r_j)}{N} \right)^2 \geq \varepsilon^2 \right\}} \right).$$

Let  $C_a = \sup_{x \geq 0} x^2 \exp \{ -a \frac{\lambda_*^2}{2T} x \}$ . Then

$$\begin{aligned} \mathcal{E}_J \left( \|Y_j\|^2 \mathbf{1}_{\{\|Y_j\| \geq \varepsilon\}} \right) &\leq \frac{\delta \lambda_*^2 C_a}{T N^2} \mathcal{E}_J \left( \left( J_{ij} - \frac{J(r_i, r_j)}{N} \right)^{-2} e^{\left\{ \frac{a}{2} N \left( J_{ij} - \frac{J(r_i, r_j)}{N} \right)^2 \right\}} \mathbf{1}_{\left\{ \frac{\delta T}{\lambda_*^2} \left( J_{ij} - \frac{J(r_i, r_j)}{N} \right)^2 \geq \varepsilon^2 \right\}} \right) \\ &\leq \frac{C_a \delta^2}{N^2 \varepsilon^2} \exp \left\{ a \frac{\|\bar{J}\|_\infty^2}{N} \right\} \mathcal{E}_J \left( \exp \{ a N J_{ij}^2 \} \right). \end{aligned}$$

Choosing  $\varepsilon = N^{-\frac{1}{4}}$ , and using hypothesis  $(H_J)$  yields the result.

## REFERENCES

- [1] S. Amari. Characteristics of random nets of analog neuron-like elements. *Syst. Man Cybernet. SMC-2*, 1972.
- [2] I. Aradi and I. Soltesz. Modulation of network behaviour by changes in variance in interneuronal properties. *The Journal of physiology*, 538(1):227, 2002.
- [3] G Ben Arous, Amir Dembo, and Alice Guionnet. Aging of spherical spin glasses. *Probability theory and related fields*, 120(1):1–67, 2001.
- [4] Gerard Ben Arous and Alice Guionnet. Langevin dynamics for sherrington-kirkpatrick spin glasses. In *Mathematical Aspects of Spin Glasses and Neural Networks*, pages 323–353. Springer, 1998.
- [5] G. Ben-Arous and A. Guionnet. Large deviations for langevin spin glass dynamics. *Probability Theory and Related Fields*, 102(4):455–509, 1995.
- [6] P.C. Bressloff. Stochastic neural field theory and the system-size expansion. *SIAM J. on Applied Mathematics*, 70:1488–1521, 2009.
- [7] P.C. Bressloff. Spatiotemporal dynamics of continuum neural fields. *Journal of Physics A: Mathematical and Theoretical*, 45:033001, 2012.
- [8] P.C. Bressloff, J.D. Cowan, M. Golubitsky, P.J. Thomas, and M.C. Wiener. Geometric visual hallucinations, euclidean symmetry and the functional architecture of striate cortex. *Phil. Trans. R. Soc. Lond. B*, 306(1407):299–330, March 2001.
- [9] MA Buice and JD Cowan. Field-theoretic approach to fluctuation effects in neural networks. *Physical Review E*, 75(5), 2007.
- [10] MA Buice, JD Cowan, and CC Chow. Systematic fluctuation expansion for neural network activity equations. *Neural computation*, 22(2):377–426, 2010.
- [11] Michael A Buice and Carson C Chow. Beyond mean field theory: statistical field theory for neural networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(03):P03003, 2013.
- [12] Michael A Buice and Carson C Chow. Dynamic finite size effects in spiking neural networks. *PLoS Comput. Biol.*, 9:e1002872, 2013.
- [13] Tanguy Cabana and Jonathan Touboul. Large deviations and dynamics of large stochastic heterogeneous neural networks. *Journal of Statistical Physics*, Volume 153, Issue 2, pp 211–269, 2013.
- [14] D Cai, L Tao, M Shelley, and DW McLaughlin. An effective kinetic representation of fluctuation-driven neuronal networks with application to simple and complex cells in visual cortex. *Proceedings of the National Academy of Sciences*, 101(20):7757–7762, 2004.
- [15] Bruno Cessac and Manuel Samuelides. From neuron to neural networks dynamics. *EPJ Special Topics in Dynamical Neural Networks*, 142(1):7–88, 2007.
- [16] G Da Prato and J Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge Univ Pr, 1992.
- [17] Paolo Dai Pra and Frank den Hollander. McKean-vlasov limit for interacting random processes in random media. *Journal of statistical physics*, 84(3–4):735–772, 1996.
- [18] E. Dauce, O. Moynot, O. Pinaud, and M. Samuelides. Mean-field theory and synchronization in random recurrent neural networks. *Neural Processing Letters*, 14(2):115–126, 2001.
- [19] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Verlag, 2009.
- [20] J.D. Deuschel and D.W. Stroock. *Large deviations*, volume 137. Academic Pr, 1989.
- [21] G Bard Ermentrout and Jack D Cowan. A mathematical theory of visual hallucination patterns. *Biological cybernetics*, 34(3):137–150, 1979.
- [22] O. Faugeras, J. Touboul, and B. Cessac. A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Neuroscience*, 3(1), 2009.
- [23] Olivier Faugeras and James Maclaurin. Asymptotic description of stochastic neural networks. i. existence of a large deviation principle elastic neural networks. ii. characterization of the limit law. *Comptes Rendus Mathématique*, 352(10):847–852, 2014.
- [24] Shintaro Funahashi, Charles J Bruce, and Patricia S Goldman-Rakic. Mnemonic coding of visual space in the monkey’s dorsolateral prefrontal cortex. *Journal of neurophysiology*, 61(2):331–349, 1989.

- [25] A. Guionnet. Averaged and quenched propagation of chaos for spin glass dynamics. *Probability Theory and Related Fields*, 109(2):183–215, 1997.
- [26] Geoffroy Hermann and Jonathan Touboul. Heterogeneous connections induce oscillations in large-scale networks. *Physical review letters*, 109(1):018702, 2012.
- [27] Dirk Jancke, Frédéric Chavane, Shmuel Naaman, and Amiram Grinvald. Imaging cortical correlates of illusion in early visual cortex. *Nature*, 428(6981):423–426, 2004.
- [28] Zachary P Kilpatrick. Interareal coupling reduces encoding variability in multi-area models of spatial working memory. *Frontiers in computational neuroscience*, 7, 2013.
- [29] Eric Luçon, Wilhelm Stannat, et al. Mean field limit for disordered diffusions with singular interactions. *The Annals of Applied Probability*, 24(5):1946–1993, 2014.
- [30] Cheng Ly and Daniel Tranchina. Critical analysis of dimension reduction by a moment closure method in a population density approach to neural network modeling. *Neural Computation*, 19(8):2032–2092, Aug 2007.
- [31] Xuerong Mao. *Stochastic Differential Equations and Applications*. Horwood publishing, 2008.
- [32] O. Moynot and M. Samuelides. Large deviations and mean-field theory for asymmetric random recurrent neural networks. *Probability Theory and Related Fields*, 123(1):41–75, 2002.
- [33] Lyle Muller, Alexandre Reynaud, Frédéric Chavane, and Alain Destexhe. The stimulus-evoked population response in visual cortex of awake monkey is a propagating wave. *Nature communications*, 5, 2014.
- [34] Aaditya V Rangan, David Cai, and Louis Tao. Numerical methods for solving moment equations in kinetic theory of neuronal network dynamics. *Journal of Computational Physics*, 221(2):781–798, 2007.
- [35] AV Rangan, G Kova i, and D Cai. Kinetic theory for neuronal networks with fast and slow excitatory conductances driven by the same spike train. *Physical Review E*, 77(4), 2008.
- [36] D Revuz and M Yor. *Continuous martingales and Brownian motion*. Springer Verlag, 1999.
- [37] H. Sompolinsky, A. Crisanti, and HJ Sommers. Chaos in Random Neural Networks. *Physical Review Letters*, 61(3):259–262, 1988.
- [38] AS Sznitman. Equations de type de boltzmann, spatialement homogenes. *Probability Theory and Related Fields*, 66(4):559–592, 1984.
- [39] AS Sznitman. Topics in propagation of chaos. *Ecole d’Eté de Probabilités de Saint-Flour XIX*, pages 165–251, 1989.
- [40] Michel Talagrand. *Mean field models for spin glasses: Volume I: Basic examples*, volume 54. Springer Science & Business Media, 2010.
- [41] Michel Talagrand. *Mean field models for spin glasses: Volume II: Advanced Replica Symmetry and Low Temperature*, volume 54. Springer Science & Business Media, 2010.
- [42] Jonathan Touboul. On the dynamics of mean-field equations for stochastic neural fields with delays. *Physica D: Nonlinear Phenomena*, 241(15):1223–1244, 2012.
- [43] Jonathan Touboul. Spatially extended networks with singular multi-scale connectivity patterns. *Journal of Statistical Physics*, 156(3):546–573, 2014.
- [44] Jonathan Touboul et al. Propagation of chaos in neural fields. *The Annals of Applied Probability*, 24(3):1298–1328, 2014.
- [45] H.R. Wilson and J.D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12:1–24, 1972.
- [46] H.R. Wilson and J.D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biological Cybernetics*, 13(2):55–80, September 1973.